Titre: Title:	Towards Sustainable Human-Robot Collaboration Disassembly Planning: Reinforcement Learning-Centered Hybrid Approaches with Fuzzy Logic
Auteur: Author:	Ashkan Amirnia
Date:	2025
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Amirnia, A. (2025). Towards Sustainable Human-Robot Collaboration Disassembly Planning: Reinforcement Learning-Centered Hybrid Approaches with Fuzzy Logic [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/69044/

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Towards Sustainable Human-Robot Collaboration Disassembly Planning: Reinforcement Learning-Centered Hybrid Approaches with Fuzzy Logic

ASHKAN AMIRNIA
Département de mathématiques et de génie industriel

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*
Génie industriel

Mai 2025

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Cette thèse intitulée :

Towards Sustainable Human-Robot Collaboration Disassembly Planning: Reinforcement Learning-Centered Hybrid Approaches with Fuzzy Logic

présentée par **Ashkan AMIRNIA**
en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Jean-Marc FRAYRET, président
Samira KEIVANPOUR, membre et directrice de recherche
Abdelhak OULMANE, membre
Bertrand LARATTE, membre externe

DEDICATION

During the trip, the dwellers of the sacred sanctuary, adorned with purity and holy modesty, have drunk the wine of love with me.

Hafez Shirazi

ACKNOWLEDGEMENTS

First, I would like to express my deepest gratitude to my dear supervisor, Dr. Samira Keivanpour, for trusting me and providing me with the opportunity to pursue my PhD under her guidance. I am thankful for her sincere help, tireless efforts, and compassionate encouragement, which greatly motivated me to continue on this path. It was an absolute pleasure to be a part of her incredible PolyCircle X.0 research team.

I am honored to have Dr. Frayret, Dr. Oulmane, and Dr. Laratte as members of my jury. I deeply appreciate their time, expertise, and insights.

During my PhD, I had the honor of collaborating alongside my friends Ahmad, Aysan, Elham, Hermes, Reda, Adel, Soumayya, Anuradha, Saeid, Sadaf, Tomas, Irune, Mahboobeh, Eli, and Hadi. They are all very valuable to me for the great experiences we shared and the lessons I learned from them.

I am also grateful to dear Francois Juteau, Zeina Tamaz, and Suzanne Guindon for their wonderful kindness.

I would like to thank my old friends Ali, Tareqh, Alireza, and Amin for their enduring support and invaluable friendships over many years.

I am grateful to my former professors, Amir Arsalan Saghaeian, Dr. Ehsan Namjoo, and Dr. Fatemeh Akbarzadeh. Beyond the academic knowledge they offered, they taught me great lessons about integrity, morality, and compassion.

Foremost, I would like to express my heartfelt appreciation to my family for their endless love. They provided me with the opportunity to discover and pursue my passions and supported me along the way. None of this would have been possible without their help. Undoubtedly, they are my most valuable treasure.

Last but not least, I would like to thank God for the hidden scenes still playing out along the way.

RÉSUMÉ

Le désassemblage des produits en fin de vie est une étape importante de la refabrication, qui se traduit par des avantages économiques et environnementaux significatifs. Traditionnellement, les opérations de désassemblage sont effectuées manuellement, ce qui est coûteux, prend du temps et nécessite des opérateurs humains qualifiés. En outre, la qualité des résultats dépend fortement des compétences des opérateurs. Le passage du désassemblage manuel au désassemblage automatisé par l'utilisation de robots collaboratifs (cobots) en tant qu'assistants humains offre un potentiel important d'amélioration de l'efficacité et de la qualité des processus. Les cobots peuvent exécuter efficacement des tâches simples et répétitives ainsi que des tâches dangereuses qui présentent des risques élevés pour la santé humaine. En revanche, les cobots ne peuvent pas exécuter des tâches difficiles et complexes qui requièrent les compétences et la force des opérateurs humains. C'est pourquoi les méthodes de désassemblage en collaboration homme-robot (HRC), qui exploitent les capacités complémentaires des humains et des cobots dans un processus en même temps, deviennent de plus en plus populaires.

Malgré tous les avantages du désassemblage HRC par rapport au désassemblage traditionnel, la recherche dans ce domaine en est encore à ses débuts. Cette thèse présente de multiples contributions pour relever les défis existants et combler les lacunes potentielles dans la littérature. Comme première contribution, nous présentons un modèle d'apprentissage par renforcement (RL) multi-agents pour la planification du désassemblage HRC qui optimise dynamiquement le processus en tenant compte de plusieurs facteurs, y compris le temps d'opération et les différentes caractéristiques des composants du produit. En utilisant une approche basée sur les graphes, le modèle représente la structure du produit et prend en compte les dépendances des tâches.

Deuxièmement, cette recherche présente un nouveau modèle durable de planification du désassemblage basé sur la logique logique (RL) qui optimise le processus en tenant compte non seulement des objectifs économiques mais aussi des objectifs environnementaux et sociaux. Dans le cas présent, les objectifs économiques sont la minimisation de la durée des opérations et de la fréquence de changement d'outil. Les objectifs environnementaux sont la minimisation de la consommation d'énergie du cobot et la maximisation de la qualité des pièces récupérées, tandis que les facteurs sociaux comprennent le risque ergonomique et la sécurité humaine. En outre, une approche floue est présentée pour modéliser les paramètres incertains de l'environnement.

Troisièmement, nous introduisons une approche basée sur la vision informatique multi-caméras et la logique floue pour prendre en compte les risques ergonomiques associés à chaque opération de démontage. Quatrièmement, nous proposons une approche de planification du désassemblage du HRC basée sur un cadre fuzzy-RL, dans lequel un modèle flou sert de copilote pour le modèle RL afin d'améliorer ses performances. Ce cadre comprend également un module collaboratif d'intelligence qui sélectionne l'un des modèles flous ou RL pour prendre des décisions à chaque étape. Enfin, cette thèse examine en détail les applications de l'IA dans le traitement des aéronefs EoL, en se concentrant sur les processus de désassemblage, de recyclage et d'entretien des produits. Elle identifie les problèmes actuels et les lacunes potentielles, en soulignant les nouvelles opportunités pour les recherches futures.

Afin d'évaluer les contributions proposées, cette thèse utilise plusieurs ensembles de données, mesures et analyses. Les critères d'évaluation comprennent les valeurs de récompense et le temps nécessaire à la convergence. En outre, nous avons effectué diverses analyses pour évaluer les performances des modèles dans différents scénarios. Il s'agit notamment d'analyses de sensibilité des paramètres les plus critiques des modèles, d'une analyse de compromis de l'importance des objectifs dans le processus décisionnel et de l'évaluation des performances des modèles dans des conditions incertaines qui contiennent des tâches avec une probabilité d'échec et des temps d'exécution variables. En outre, nous avons développé une interface utilisateur graphique (GUI) qui permet aux utilisateurs de personnaliser le processus en ajustant l'importance des objectifs de durabilité. En outre, un cadre expérimental a été mis en place pour valider l'approche d'évaluation des risques ergonomiques proposée.

ABSTRACT

End-of-life (EoL) product disassembly is an important step in remanufacturing, resulting in significant economic and environmental benefits. Traditionally, disassembly operations are carried out manually, which is costly, time-consuming, and requires skilled human operators. Furthermore, output quality highly relies on operators' skills. Transitioning from manual disassembly to automated disassembly by employing collaborative robots (cobots) as human assistants has significant potential to improve process efficiency and quality. Cobots can effectively perform simple and repetitive tasks as well as dangerous tasks that pose high risks to human health. In contrast, cobots cannot execute challenging and complex tasks that require the skill and strength of human operators. Therefore, human-robot collaboration (HRC) disassembly methods, which exploit the complementary capabilities of humans and cobots in a process at the same time, are becoming more and more popular.

Despite all the advantages of HRC disassembly over traditional disassembly, research in this domain is still in its early stages. This thesis presents multiple contributions to address the existing challenges and fill the potential gaps in the literature. As a first contribution, we present a multi-agent reinforcement learning (RL) model for HRC disassembly planning that dynamically optimizes the process by considering several factors, including operation time and different characteristics of product components. Using a graph-based approach, the model represents the product structure and considers task dependencies.

Secondly, this research presents a novel sustainable RL-based disassembly planning model that optimizes the process by addressing not only economic objectives but also environmental and social objectives. In this case, it considers minimizing operation time and tool change frequency as the economic objectives. The environmental objectives are minimizing cobot energy consumption and maximizing recovered parts quality, while the social factors include ergonomic risk and human safety. Furthermore, a fuzzy-based approach is presented to model uncertain parameters in the environment.

Thirdly, we introduce an approach based on multi-camera computer vision and fuzzy logic to consider the ergonomic risks associated with each disassembly operation. Fourthly, we propose an HRC disassembly planning approach based on a fuzzy-RL framework, in which a fuzzy model serves as a copilot for the RL model to improve its performance. This framework also includes an intelligence collaborative module that selects one of the fuzzy or RL models for making decisions at every step. Finally, this thesis comprehensively reviews AI applications in EoL aircraft treatment, focusing on product disassembly, recycling, and maintenance

processes. It identifies current issues and potential gaps, highlighting new opportunities for future investigations.

In order to evaluate the proposed contributions, this thesis uses multiple datasets, metrics, and analyses. The evaluation criteria include reward values and the time required for convergence. Furthermore, we conducted various analyses to assess the models' performance under different scenarios. They include sensitivity analyses of the most critical parameters of the models, a trade-off analysis of objectives' importance in the decision-making process, and the evaluation of the models' performance under uncertain conditions that contain tasks with a probability of failure and varying execution times. In addition, we developed a graphical user interface (GUI) that enables users to customize the process by adjusting the importance of sustainability objectives. Furthermore, an experimental setting has been conducted to validate the proposed ergonomic risk assessment approach.

TABLE OF CONTENTS

DEDICATION iii

ACKNOWLEDGEMENTS iv

RÉSUMÉ v

ABSTRACT vii

TABLE OF CONTENTS ix

LIST OF TABLES xiii

LIST OF FIGURES xiv

LIST OF SYMBOLS AND ACRONYMS xvi

CHAPTER 1 INTRODUCTION 1
1.1 Challenges and gaps 2
1.2 Research objectives and contributions 4
1.3 Publications 6
1.4 Thesis outline 7

CHAPTER 2 LITERATURE REVIEW 8
2.1 Classical optimization algorithms in disassembly planning 8
2.2 AI-driven approaches in product disassembly 10
2.2.1 Classical machine learning 10
2.2.2 Reinforcement learning 10
2.2.3 Computer vision 11
2.3 Energy consumption in disassembly processes 12
2.4 Human safety in disassembly processes 13
2.5 Ergonomic risk in disassembly processes 13
2.6 Technical feasibility in disassembly processes 14
2.7 Design for disassembly 15
2.8 Human-robot interactions 16
2.9 Joint applications of fuzzy logic and ML 19

CHAPTER 3 METHODOLOGY 21
3.1 First article 22
3.2 Second article 24
3.3 Third article 25
3.4 Fourth article 26

CHAPTER 4 ARTICLE 1: A CONTEXT-AWARE REAL-TIME HUMAN-ROBOT COLLABORATING REINFORCEMENT LEARNING-BASED DISASSEMBLY PLANNING MODEL UNDER UNCERTAINTY 29
4.1 Abstract 29
4.2 Introduction 29
4.3 Background 33
4.3.1 Disassembly processes 33
4.3.2 Human-robot interactions in disassembly planning 35
4.3.3 Deep reinforcement learning in manufacturing 37
4.3.4 The required disassembly features comparison and the potential gaps 38
4.4 Methodology 39
4.4.1 A product architecture modeling 40
4.4.2 Task planning strategy based on the disassembly difficulty 42
4.4.3 Modeling a collaborative disassembly problem based on reinforcement learning 43
4.4.4 Switching between the collaborative and inactive modes 45
4.4.5 Real-time task allocation process 46
4.4.6 A multi-agent deep reinforcement learning approach for disassembly planning 47
4.5 Experiments 49
4.5.1 Experimental setup 51
4.5.2 Results and discussion 53
4.6 Conclusion 60

CHAPTER 5 ARTICLE 2: REAL-TIME SUSTAINABLE COBOTIC DISASSEMBLY PLANNING USING FUZZY REINFORCEMENT LEARNING 63
5.1 Abstract 63
5.2 Introduction 63
5.3 Related works 66
5.3.1 Disassembly planning 66
5.3.2 Human-robot collaboration 67

5.3.3 Integrating sustainability criteria into disassembly planning 68
5.4 Problem statement 73
5.5 Methodology 76
5.5.1 Proposed RL-based sustainable HRC disassembly planning model . . 78
5.5.2 Fuzzy-based disassembly environment 82
5.5.3 The proposed model architecture 84
5.6 Experiments 86
5.6.1 Experimental setup 86
5.6.2 Results and discussion 88
5.7 Conclusion 95

CHAPTER 6 ARTICLE 3: REAL-TIME VIDEO PROCESSING IN FUZZY POSTURE-BASED ERGONOMIC ANALYSIS IN A DISASSEMBLY CELL 101
6.1 Abstract 101
6.2 Introduction 101
6.3 Problem context 103
6.3.1 Product disassembly 103
6.3.2 Ergonomic consideration in disassembly cells 104
6.3.3 Research gaps in ergonomic assessment tools 105
6.4 Proposed fuzzy approach 105
6.4.1 Fuzzy ergonomic assessment tool 106
6.4.2 Real-time video processing method 107
6.5 Application perspective 108
6.6 Conclusion 109

CHAPTER 7 ARTICLE 4: A FUZZY-GUIDED REINFORCEMENT LEARNING METHOD FOR SUSTAINABLE COBOTIC DISASSEMBLY PLANNING UNDER UNCERTAINTY 111
7.1 Abstract 111
7.2 Introduction 111
7.3 Related works 114
7.3.1 Classical planning models 115
7.3.2 Learning-based planning models 116
7.3.3 Comparison of the proposed approach with existing methods 117
7.4 Problem statement 119
7.5 Methodology 121
7.5.1 Main architecture 121

7.5.2 RL models . . . 123
7.5.3 Fuzzy models . . . 127
7.5.4 Environment modeling . . . 129
7.5.5 Failure operations probability . . . 129
7.5.6 The multi-agent fuzzy-RL algorithm for HRC disassembly planning . 131
7.6 Results and discussion . . . 131
7.7 Conclusion . . . 138

CHAPTER 8 CONCLUSION AND GENERAL DISCUSSION . . . 144
8.1 The first article . . . 144
8.2 The second article . . . 145
8.3 The third article . . . 145
8.4 The fourth article . . . 146
8.5 Implications . . . 147
8.6 Limitations and future research . . . 148

REFERENCES . . . 149

APPENDICES . . . 163

LIST OF TABLES

Table 4.1	Comparison of required disassembly features in the various models	39
Table 4.2	Operators' types and their proper tasks in terms of disassembly difficulty	43
Table 4.3	Description list of key symbols	49
Table 4.4	Disassembly features and their corresponding values	52
Table 4.5	Time and recovered part quality of disassembly tasks in each difficulty level based on operators' skill level	55
Table 5.1	A qualitative analysis of the proposed model with the recent literature	72
Table 5.2	Parameters in the proposed fuzzy-RL model	83
Table 5.3	The utilised case study and its corresponding features values	87
Table 6.1	Posture risk of upper arm (shoulder)	106
Table 6.2	Posture risk of lower arm (elbow)	107
Table 6.3	Integrating fatigue parameter to posture risk.	107
Table 6.4	Posture risks based on angles detected for disassembly tasks.	110
Table 6.5	Final and cumulative risks of each body part.	110
Table 7.1	A qualitative comparison of the proposed approach with the recent literature	118
Table 7.2	Cobot fuzzy rules	128
Table 7.3	Human fuzzy rules	128
Table 7.4	Time, recovered quality, and penalty in various conditions for the cobot operator	129
Table 7.5	Time, recovered quality, and penalty in various conditions for the human operator	130
Table 7.6	The model configurations based on the impactful factors in the exploitation/exploration trade-off	134
Table 7.7	The case study feature values	143

LIST OF FIGURES

Figure 1.1	The objectives and contributions of thesis	6
Figure 2.1	The outline of the literature review	9
Figure 2.2	Human-robot interactions categories	17
Figure 2.3	A human-robot collaboration cell	18
Figure 3.1	The presented articles and their interconnections	21
Figure 3.2	An example of the real-time task allocation mechanism	22
Figure 3.3	A graph of an EoL product	23
Figure 3.4	The proposed multi-agent RL-based model	23
Figure 3.5	The proposed multi-agent RL-based framework	24
Figure 3.6	The fuzzy-based environment	25
Figure 3.7	Ergonomic risks corresponding to disassembly activities	26
Figure 3.8	The proposed approach for ergonomic risk assessment	26
Figure 3.9	The multi-agent fuzzy-RL framework	27
Figure 3.10	The developed GUI	28
Figure 4.1	Human-robot interaction categories	36
Figure 4.2	A graph of an EoL product	40
Figure 4.3	Switching between the collaborative and inactive modes	46
Figure 4.4	The real-time task allocation mechanism	47
Figure 4.5	The proposed reinforcement learning-based framework	49
Figure 4.6	The product structure presented in [1]	51
Figure 4.7	Sensitivity analyses in Scenario I	56
Figure 4.8	Sensitivity analyses in Scenario II	58
Figure 4.9	Sensitivity analyses in Scenario III	59
Figure 4.10	A qualitative comparison of recent literature	61
Figure 5.1	The graph representation approach	77
Figure 5.2	Sustainable development disassembly planning triangle	78
Figure 5.3	The proposed fuzzy environment	83
Figure 5.4	Architecture of the sustainable HRC disassembly framework using fuzzy-RL	85
Figure 5.5	The structure of the product introduced in [1]	86
Figure 5.6	Sensitivity analysis in Scenario I	90
Figure 5.7	Sensitivity analysis in Scenario II	91
Figure 5.8	Sensitivity analysis in Scenario III	92

Figure 5.9 The reward values for sensitivity analysis on decision-making under uncertain conditions 94
Figure 5.10 The required iterations for convergence for the proposed model and the baseline paper on a single dataset 95
Figure 5.11 The employed fuzzy numbers in the proposed model 99
Figure 6.1 The proposed vision-based system for analyzing ergonomic risks . . . 106
Figure 6.2 Two vectors and the corresponding angle between them 108
Figure 6.3 Detected joints and computed angles in the utilised case study. The collected image data was used for initial feasibility testing and was not subject to detailed analysis in this study 109
Figure 7.1 The conceptual framework for fuzzy-RL integration 121
Figure 7.2 The proposed hybrid fuzzy-RL model 122
Figure 7.3 The multi-agent fuzzy-RL framework 123
Figure 7.4 The state vectors of the human and cobot agents 125
Figure 7.5 The graph representation approach 126
Figure 7.6 Effectiveness fuzzy membership 128
Figure 7.7 The flowchart of the proposed HRC disassembly planning algorithm 132
Figure 7.8 Reward values resulted from different configurations 135
Figure 7.9 Sensitivity analysis on real-time decision-making 137
Figure 7.10 Sensitivity analysis on sustainable objectives importance 137
Figure 7.11 The developed GUI 139
Figure 7.12 Sensitivity analysis on operation failure 140

LIST OF SYMBOLS AND ACRONYMS

EoL	End of life
HRC	Human-robot collaboration
AI	Artificial intelligence
RL	Reinforcement learning
DQN	Deep Q-network
MDP	Markov decision process
LSTM	Long short-term memory
GA	Genetic algorithm
PSO	Particle swarm optimization
ACO	Ant colony optimization
ABC	Artificial bee colony
SVM	Support vector machine
DCG	Disassembly constraint graph
YOLO	You only look once
MTTR	Mean time to repair
DfD	Design for disassembly
MOST	Maynard operation sequence technique
eDiM	Ease of disassembly metrics
EPN	Extended petri net
VAE	Variational autoencoder
HSV	Hue, saturation, and value
FRL	Fuzzy reinforcement learning
CNN	Convolutional neural network
SCIP	Solving constraint integer programs
WEEE	Waste electrical and electronic equipment
CAD	Computer-aided design
CT	Cycle time
REBA	Rapid entire body assessment
RULA	Rapid upper limb assessment
EATs	Ergonomic assessment tools
SAE	Stacked autoencoder
SQP	Sequential quadratic programming
GUI	Graphical user interface

WMSDs	Work-related musculoskeletal disorders
VR	Virtual reality
XR	Extended reality
ML	Machine learning
SG	Straight grain
WG	Wavy grain
WS	Wood stain
CD	Color difference
FFT	Fast Fourier transform
HMM	Hidden Markov model
HitL	Human-in-the-loop
HotL	Human-on-the-loop
EV	Electric vehicle
KPIs	Key performance indicators
FIS	Fuzzy-inference system
MVC	Maximum voluntary contraction
DDPG	Deep deterministic policy gradient
XRF	X-ray fluorescence
SIFT	Scale-invariant feature transform
MIS	Magnetic induction spectroscopy
ANN	Artificial neural network
ROI	Region of interest
LIBS	Laser-induced breakdown spectroscopy
LCA	Life cycle assessment

CHAPTER 1 INTRODUCTION

Nowadays, population growth, rising consumerism, and the shift towards modernity have significantly increased the production of goods. This has resulted in the generation of vast amounts of end-of-life (EoL) products each year, posing serious environmental concerns. Landfilling, incineration, recycling, and remanufacturing are among the most common solutions to address this challenge. However, landfilling or incineration pollutes the air, soil, and water. While recycling is a more eco-friendly solution, it may consume considerable amounts of energy. On the other hand, remanufacturing focuses on rebuilding EoL products using various methods instead of discarding, burning, or recycling them [2]. A key step in remanufacturing is product disassembly, which refers to the process of separating a product into its components. This procedure offers significant environmental and economic benefits, such as waste reduction, lower costs of waste management, preserving natural resources and raw materials, and energy footprint mitigation.

Conventionally, disassembly operations are manually executed by labours, a time-intensive and expensive approach. It also exposes humans to executing dangerous and unsafe tasks. In addition, manual disassembly processes pose other challenges, such as complex product designs, component damage, low efficiency, and waste generation, which impact the feasibility and efficiency of remanufacturing. Over the past decades, the manufacturing industry has experienced significant progress. In this way, manufacturing processes, such as disassembly, are switched from manually operated procedures to automatic and autonomous operations. This transition aims to efficiently and effectively reduce costs and increase profits. Similarly, the use of collaborative robots (cobots) in disassembly processes is becoming more and more popular, and these processes are being transformed through collaboration between humans and cobots. Cobots carry out disassembly tasks in a cell alongside one or more human operators. Cobots are able to perform tasks that may harm humans or have ergonomic risks for them. These tasks include different activities dealing with hazardous substances, exposure to toxic gases, or performing repetitive operations, which may negatively affect the spine and skeleton over time, increasing the potential for work-related musculoskeletal disorders (WMSDs) [3]. Moreover, performing simple and repetitive tasks or working for prolonged hours may reduce human awareness and cause human fatigue, resulting in a human operator's poor performance. In contrast, cobots can complete such tasks with high precision. However, the major limitation of the cobots is their insufficient flexibility and strength to complete complex and difficult tasks. On the other hand, human operators can effectively execute such complex and heavy tasks. Therefore, human-robot collaboration (HRC) leverages both

the precision of cobots as well as the flexibility of human operators, promising significant accomplishments in this context.

Although involving cobots in such processes may enhance efficiency and quality, it is crucial to consider the amount of energy consumed by cobots. The increased use of cobots causes more energy consumption, which not only increases costs but also generates excessive energy-related pollutants. Furthermore, performing complex and heavy tasks with cobots could potentially damage the product or cobots physical structures. It is also critical to consider human factors such as ergonomic risks and human safety while allocating tasks to operators. Thus, optimized task assignment is significantly important as inefficient allocating tasks may greatly reduce process productivity. Disassembly planning refers to sustainably optimizing the disassembly process to not only maximize profits but also to minimize operation time, the utilization of raw materials, harmful environmental factors, such as CO2 emissions, and other factors influencing costs. A disassembly plan includes three main steps: preprocessing, modeling, and sequential planning. The preprocessing step involves all the preliminary activities required to start the process. These activities include collecting information about the product, its corresponding components and connections, and other related data obtained from a Computer-aided design (CAD) model or other tools. Combining different mathematical approaches, the disassembly modeling step aims to represent an EoL product's architecture during the disassembly process. Graphs, matrices, Petri networks, and universal techniques are the four main classes to model an EoL product [2]. In addition, sequential disassembly planning addresses the utilization of different data-driven techniques to determine the optimal task sequence for each operator.

HRC disassembly processes encompass many uncertain factors. One example is the variable performance of human operators in executing tasks due to fatigue and distractions. Furthermore, the unstable conditions of EoL products are another uncertain factor, resulting from various elements such as damage, tear, and wear of parts due to different usage patterns. These uncertainties in addition to other uncertain factors such as machine failures, human errors, damaged tools, and any external disruptions may result in the prolongation or even failure of operations. Disassembly planning models should be capable of addressing these uncertainties during the task allocation process.

1.1 Challenges and gaps

Despite all the advantages of HRC disassembly, the planning models proposed in the literature are still at the early stage. Synthesis in the literature reveals the following gaps:

1. Most existing models are unable to make decisions in real time. They statically generate a preplanned task sequence for each operator based on theoretical assumptions. These models cannot cope with uncertain conditions that may diverge a process from its expected direction, which limits their feasibility in real-world scenarios. Herein, if all decisions are specified in advance, an operation may fail. Therefore, real-time decision-making is necessary to respond quickly to changes during execution.

2. Human operators have different skill levels, affected by factors, such as experience, precision at work, and physical characteristics like agility, height, and weight. Performing disassembly tasks by humans is a skill-based activity, in which an operator's expertise impacts output quality. However, accounting for human skill levels in disassembly planning is not adequately addressed in the literature and could be explored more by researchers.

3. EoL products are composed of components with multiple features, such as weight, size, and volume. These features contribute to varying levels of task difficulty, complexity, and execution time. Although it is necessary to include these features in disassembly planning, most models in the literature do not consider them.

4. It is crucial to consider social factors such as human safety and ergonomic risk in disassembly processes. It is also essential to optimize the process with respect to the amount of energy consumed by cobots and the quality of recovered parts, which are environmental factors. Unfortunately, most previous studies only focused on economic objectives and did not address environmental and social factors in the planning process.

5. As previously mentioned, disassembly processes involve uncertainties that may lead to different outcomes when the same operator performs a given task at differing times. A disassembly planning model should be able to effectively handle this lack of certainty in data. Unfortunately, most existing approaches in the literature have been developed and validated under controlled and certain scenarios.

6. In a disassembly process, a cobot may fail to complete some tasks because of abnormal conditions, including machine failure, wear and tear tools, cobot inefficiency, or any other external interruption, which direct the task sequence planning process. In such cases, the model should reconfigure the task sequence from its original outline. Therefore, it is crucial to consider task failure in the modeling process to ensure the model's effectiveness in industrial settings. However, probability of the operation failing is not well-addressed in the literature.

7. A common issue with previous models is that users cannot easily work with and customize them. The target group in this context is manufacturers who may not be experts in computer science or data-driven models. Therefore, developing a user-friendly interface that enables manufacturers to customize models based on their specific requirements makes studies more practical for industrial implementations.

1.2 Research objectives and contributions

This thesis aims to address the mentioned gaps by presenting four research articles based on four contributions. As illustrated in Figure 1.1, the first article aims to effectively fill gaps 1, 2, and 3. Furthermore, the second article is concerned with gaps 4 and 5. The third article presents a novel approach for ergonomic risk assessment in disassembly operations, addressing part of the social factors identified in gap 4. Lastly, the fourth article addresses gaps 6 and 7. In the following, we briefly introduce these articles.

1. We present a graph-based multi-agent RL model for HRC disassembly planning. The model contains a human agent and a cobot agent, each selecting tasks for its corresponding operator. Instead of using a pre-planned task sequence, the model makes real-time decisions, enabling it to cope with the uncertain nature of the problem. It also incorporates several influential factors, such as human operators' skill level, recovered parts quality, and various parts' features, into the planning process. By using a graph-based approach, the model effectively represents a product's architecture.

2. This contribution presents a sustainable data-driven model based on RL for HRC disassembly planning. In general, the term *sustainability* refers to a broad concept encompassing environmental, social, economic, and governance aspects across various settings. This study focuses on the triple-bottom line model of sustainability, including environmental, social, and economic dimensions. In addition, it addresses only operational sustainability elements, without considering potential second-order or rebound effects, as they fall outside the scope of this research. The sustainability indicators used to assess each pillar are defined as follows:

   (a) Environmental: Cobot energy consumption and recovered parts quality

   (b) Social: Ergonomic risk and human safety

   (c) Economic: Operational time and tool change frequency.

   In this regard, the model aims to maximize recovered parts quality and human safety, while minimizing energy consumption, ergonomic risks, operation times, and tool change

frequencies. Furthermore, to handle the lack of certainties in data, we represent uncertain parameters, such as difficulty and recovered part quality, through an approach based on fuzzy logic.

3. In line with the shift toward sustainable product disassembly, this contribution introduces a novel approach based on image processing and fuzzy logic to assess ergonomic risks associated with disassembly operations. The approach leverages a multi-camera vision configuration that consists of three cameras, installed at the front and on both sides of the disassembly site. In each frame of these videos, we detect the upper body joints of a human operator engaged in a disassembly process and then compute the angles between the joints. This research uses three cameras as it is impossible to calculate all angles from videos captured with a single camera. Next, the developed fuzzy model determines the ergonomic risk—low, medium, or high—for each upper body part during an operation based on the corresponding joint angles and the time required for completing the operation. As a result, this study assesses the ergonomic risk posed to each upper body part in every operation.

4. This study develops a novel multi-agent framework based on fuzzy logic and RL for HRC disassembly planning. This framework consists of two agents, human and cobot, each using an RL model and a fuzzy model. At each time-step, one of these models is chosen to select a component for the respective operator. Since RL models often perform poorly in the initial steps, a fuzzy model is involved in each agent to enhance its corresponding RL model's performance. In the early steps, a fuzzy model in each agent is chosen more frequently for component selection, and as the process evolves, the fuzzy model is selected less and less. Through the recursive learning process, the RL model dynamically learns from the results obtained by the fuzzy model. In this way, the fuzzy model plays the role of a copilot for the RL model during the initial steps. In addition, as the fuzzy models are developed logically and based on expert knowledge, we transfer human experts' knowledge to the RL models through this framework. Herein, we use the epsilon-greedy algorithm to select between the RL and fuzzy models. If epsilon is greater than a value of N, sampled from a Gaussian distribution, the fuzzy model is chosen; otherwise, the RL model is selected. Furthermore, we diminish the epsilon value by a factor at the end of each episode. Also, as a novelty, this research conducted multiple sensitivity analyses to evaluate the framework's performance in uncertain scenarios, such as task failure probability. Additionally, we developed a graphical user interface (GUI) that enables users to easily customize the planning of disassembly processes according to required sustainability objectives.

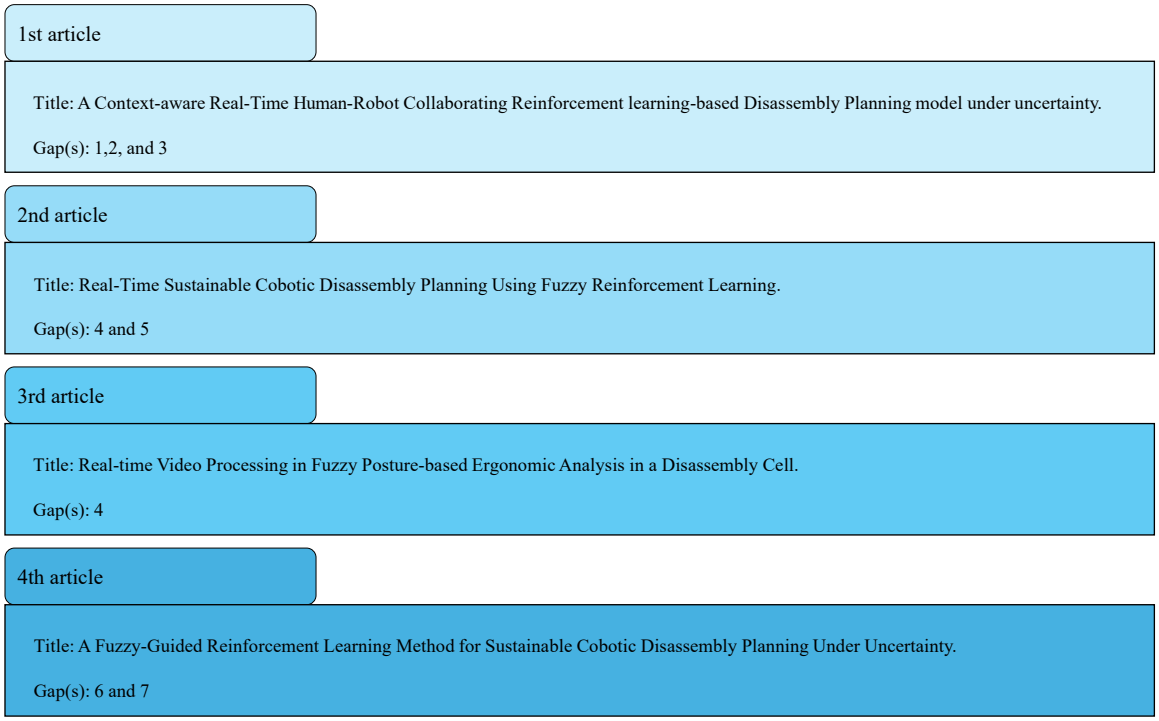


Figure 1.1 The objectives and contributions of thesis

1.3 Publications

On the basis of the contributions presented in this thesis, articles have been published or submitted, which are listed below.

- Amirnia, A., Keivanpour, S. (2024). A context-aware real-time human-robot collaborating reinforcement learning-based disassembly planning model under uncertainty. International Journal of Production Research, 62(11), 3972-3993.
- Amirnia, A., Keivanpour, S. Real-Time Sustainable Cobotic Disassembly Planning Using Fuzzy Reinforcement Learning. Accepted by International Journal of Production Research.
- Amirnia, A., Ghorbani, E., Keivanpour, S. (2024, July). Real-Time Video Processing in Fuzzy Posture-Based Ergonomic Analysis in a Disassembly Cell. In International Conference on Intelligent and Fuzzy Systems (pp. 247-256). Cham: Springer Nature Switzerland.
- Amirnia, A., Keivanpour, S. A Fuzzy-Guided Reinforcement Learning Method for Sustainable Cobotic Disassembly Planning Under Uncertainty. Submitted to International

Journal of Production Research.

Moreover, the following papers are part of my PhD research, however, they are not included in this thesis.

- Amirnia, A., Keivanpour, S. (2024, July). AI-Driven EoL Aircraft Treatment: A Research Perspective. In Intelligent Systems Conference (pp. 371-391). Cham: Springer Nature Switzerland.
- Bushehri, A. S., Amirnia, A., Belkhiri, A., Keivanpour, S., De Magalhaes, F. G., Nicolescu, G. (2023). Deep Learning-Driven Anomaly Detection for Green IoT Edge Networks. IEEE Transactions on Green Communications and Networking.
- Guillaume Mozian, Ashkan Amirnia , and Samira Keivanpour. Sustainability-aware Computer Vision for Scrap Material Recognition in Automated Sorting. Submitted to the 13th International Conference on Control, Mechatronics and Automation (ICCMA 2025).
- Saeid Jamshidi, Ashkan Amirnia, Amin Nikanjam, Kawser Wazed Nafi, Foutse Khomh, Samira Keivanpour. Self-Adaptive Cyber Defense for Sustainable IoT: A DRL-Based IDS Optimizing Security and Energy Efficiency. Submitted to Journal of Network and Computer Applications.
- Elham Haji Sami, Ahmad Shahnejat Bushehri, Ashkan Amirnia, Asad Yarahmadi, Samira Keivanpour. Integrated Sequential Matching and Routing Approach for Efficient and Eco-Friendly Freight Logistics. Submitted to Transportation Research Part C.

1.4 Thesis outline

The rest of the thesis is organized as follows. Chapter 2 reviews the literature, addressing basic concepts, modeling approaches, various data-driven planning methods, optimization objectives, and other key aspects of the problem. Chapter 3 outlines the proposed methodology to accomplish the thesis objectives. Chapters 4, 5, 6, and 7 present the research articles produced from this research. At the end, Chapter 8 concludes the thesis.

CHAPTER 2 LITERATURE REVIEW

This chapter provides a comprehensive review of the literature, as outlined in Figure 2.1. It first discusses the classical optimization algorithms in disassembly planning. Then, it addresses AI-driven approaches, including classical ML, RL, and computer vision methods, in disassembly processes. Following that, it explains the concept of design for disassembly (DfD). Next, human-robot interactions are discussed. Lastly, it reviews the joint application of fuzzy logic and ML. Although the last topic may not be directly related to other studies on disassembly and human-robot interactions, this research presents new methods based on fuzzy logic and RL for disassembly planning in HRC. Hence, previous studies on hybrid fuzzy-RL models are also discussed in this literature review.

2.1 Classical optimization algorithms in disassembly planning

There is a growing body of literature exploring how to optimize a disassembly process by sequential task allocation approaches. Researchers have widely used various rule-based and learning-based algorithms to generate optimum task sequence. Classical optimization algorithms such as Genetic algorithm (GA) [4–15], Particle swarm optimization (PSO) [16–23], and Ant colony optimization (ACO) [24–32] have been widely used in the literature. In [33], the authors proposed a hybrid optimization method as a disassembly planning model that combines the GA and the Tabu search. It also utilises a graph-based method called disassembly constraint graph (DCG) to represent a product's structure. In another study [34], McGovern and Gupta introduced an approach based on the GA for disassembly line balancing. Next, a disassembly sequence planning approach based on ACO algorithm is presented in [31]. The authors formulated the problem using three factors: the number of components, disassembly tools, and disassembly process directions. [16] introduces a disassembly planning approach that incorporates Dijkstra's algorithm and PSO, considering both functional components and fasteners.

In a recent study [35], Guo et al. developed an approach based on an improved artificial bee colony (ABC) algorithm for balancing partial multi-robotic disassembly lines. The proposed approach considers profitability, cycle time, energy consumption, and the extra time and cost resulting from workstations' reconfiguration. Fang et al. [36] proposed a multi-objective approach to balance position-constrained HRC disassembly lines. They presented a mixed integer programming model and a multi-fidelity optimization algorithm to address both small-scale and large-scale problems, respectively. Hu et al. [37] have presented a hybrid approach

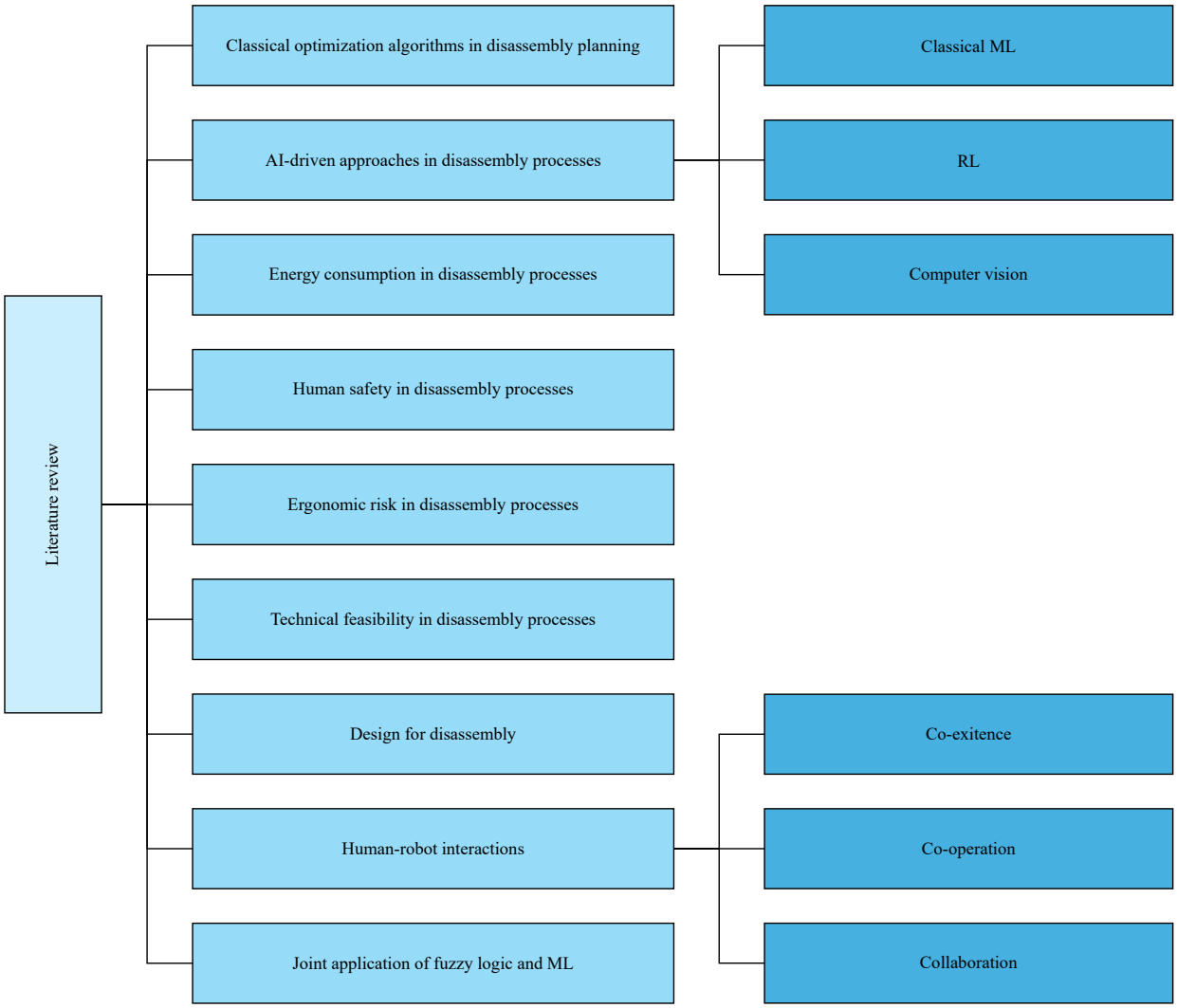


Figure 2.1 The outline of the literature review

based on general ontology and rule-based reasoning for HRC disassembly planning. They validated the proposed approach by using a gearbox as a case study. In a more recent research study [38], Lou et al. have developed a human-cyber-physical system framework for HRC disassembly planning, in which the role of humans is considered in two human-in-the-loop (HitL) and human-on-the-loop (HotL) paradigms. The authors also computed task complexity and operator ergonomics using a cloud-based approach. Finally, an enhanced hybrid grey wolf optimization algorithm plans the sequential task allocation process. They conducted an experimental analysis by applying the proposed framework to a control box as a real case study. In [39], the authors optimized a robotic disassembly planning problem with a classic multi-objective robust approach that aims to cope with uncertainties regarding products and operator conditions during a process.

2.2 AI-driven approaches in product disassembly

This section reviews AI applications in the field of product disassembly. It briefly discusses classical machine learning applications in planning. It then addresses RL models in planning followed by computer vision applications.

2.2.1 Classical machine learning

Some preliminary researchers have developed disassembly planning models using classical machine learning approaches. In addition, a combination of a Petri network and a hybrid Bayesian network is introduced as a disassembly planning model in [40]. The Petri network represents the entire process, while the hybrid Bayesian network selects the next operation in the process. Similarly, Godichaud et al. [41] presented a disassembly planning approach based on a Petri network and a Bayesian network. Xiao et al. [42] presented a method for planning electric vehicle (EV) battery disassembly. It consists of a Bayesian network and a hidden Markov model (HMM), using the forward-backward and Viterbi decoding algorithms.

2.2.2 Reinforcement learning

As discussed in Section 1, disassembly processes involve inherent uncertainties. Supervised and unsupervised ML methods generally have limited capabilities to optimize such processes. In contrast, RL models offer significant potential, as they can make decisions under uncertain scenarios while adapting to changes in surrounding conditions. There has been growing attention to utilizing RL models for the disassembly task allocation problems over the past few years. The state for these RL models is typically defined as the product's situation at each step. Moreover, the action is determined as the next component for assembly/disassembly. An early investigation of using RL for the purpose of disassembly planning was performed in [43], in which the authors presented an RL-based approach that approximates Q-values through an Eleman network. Reveliotis [44] proposed a disassembly planning method based on a dynamic programming approach that handles the uncertainty of the process as well as models the entire process through a Petri network. Mao et al. [45] have introduced an enhanced deep Q-network (DQN) model that utilises the GA to improve the long-term reward. A Petri network is also used to represent the product's architecture. The authors evaluated the proposed model for maintenance training in a virtual reality (VR) environment. Next, Chen et al. [46] have introduced a matrix-based Q-learning model to optimize a disassembly process regarding operation time. In addition, physical constraints of parts are involved in the reward function. The authors deterministically defined a time for a part disassembly as

a summation of basic disassembly time, tool changing time, the number of connections to other parts, required time for tool positioning, and cleanup time.

[47] represents disassembly process by using an extended Petri network (EPN), including features of parts, such as physical constraints, revenue, and cost. As a feature for each part in EPN, the authors considered four EoL operation choices: reuse, remanufacturing, recycling, and disposal. Next, they estimated uncertain disassembly parameters, such as revenue and cost, based on historical data by a maximum likelihood approach. They then formulated the problem into an RL framework with a tabular Q-learning algorithm. While the state is defined based on the EPN's token, the reward function comprises the difference between revenue and cost. In [48], the authors developed an RL-based disassembly planning model using a new graph-based representation approach that obtains the precedences between product's components and considers the uncertainty of an EoL product. They have presented an algorithm to assign a level to each component based on its positioning within the hierarchical structure. Accordingly, the task allocation process begins with the components at the outermost level. After disassembling all the components at this level, the algorithm moves to the next inner level. In this way, the process continues until all the components have been disassembled and the innermost level is reached. In this case, a reward function contains processing time, profit, and the difference in the levels of the previous and newly selected components. The last term awards the agent once it completes the disassembly of components in a level and moves to the next level.

An RL-based assembly planning model has been introduced in [49]. The state includes two terms: a vector representing the required assembly tool for each task and a binary vector indicating whether each task is completed. Moreover, the reward signal consists of completion time and user satisfaction regarding the assembly sequence. Herein, each component's disassembly completion time can be configured either deterministically or stochastically. In their proposed framework, the authors compared several RL models using airplane toy data [50,51] as a case study. Recently, Sadeghi Tabar et al. [52] proposed a disassembly planning approach that models the process using a Petri network and optimizes it through a Q-learning method. The authors considered time, quality, and process capability as the main objectives in the reward function. In addition, they evaluated the model with an engine starter motor as a case study.

2.2.3 Computer vision

Researchers have become interested in the applications of computer vision and image processing in disassembly processes. In [53], the authors have proposed an image processing-based

approach for screw detection in EV motors during disassembly operations. This approach integrates several image processing techniques, including the hue, saturation, value (HSV) color modeling, image depth detection, grayscale conversion, and the Harris detector [54]. In [55], the authors introduced a method using Tiny-Yolo v2 to detect screws in robotic disassembly processes.

Zhang et al. [56] presented a hybrid model based on a variational autoencoder (VAE), an HMM, and a support vector machine (SVM) for human activity recognition in HRC disassembly. In [57], the authors introduced a tool recommendation system based on computer vision methods for robotic disassembly. First, a YOLOv4 algorithm detects screws in images. Then, an EfficientNetv2 algorithm classifies the screws' shapes and recommends a tool to a robot operator.

2.3 Energy consumption in disassembly processes

Sustained economic growth relies on energy, yet several factors, such as the increasing energy demand caused by global population growth, threaten modern economic progress. The amount of consumed energy also plays a critical role in cobotic assembly/disassembly operations. Consuming more energy produces more energy footprints, which negatively impacts the environment. Moreover, uncontrolled energy consumption increases economic costs [2,58].

A few recent studies have paved the way to integrate consuming energy into the robotic disassembly planning process. [59] develops a robotic parallel disassembly sequence planning framework using the ABC algorithm, aiming to minimize the makespan and total energy consumption of robots. They assumed the consumed energy arises in three stages: during disassembling, while waiting for new tasks, and during tool change. Similarly, [60] optimizes a robotic disassembly process to achieve maximum profit while minimizing energy footprint by employing the Bees algorithm. In the context of assembly/disassembly line balancing, [61] proposes a model based on an improved bi-objective evolutionary algorithm to optimize a robotization assembly line. As two main objectives, the model attempts to minimize total energy consumption and the number of workstations. In another study, [62] introduces a modified GA algorithm to solve the assembly line balancing and part feeding problem. The algorithm considers reducing energy consumption, the number of stations, and the number of supermarkets as the main objectives. [63] proposes a mixed model to minimize cycle time, total consumed energy, and peak workstation energy consumption in a robotic disassembly line balancing problem.

Despite the mentioned considerable efforts made in the planning of robotic disassembly,

the aspect of energy in cobotic disassembly planning still remains largely unexplored by researchers. Hence, optimizing energy consumption of cobots in the disassembly process is an emerging research area.

2.4 Human safety in disassembly processes

Human safety is a crucial concern for manufacturing companies. It is essential to avoid assigning dangerous operations to humans in assembly/disassembly processes. Many EoL products, such as waste electrical and electronic equipment (WEEE), include hazardous substances, which adversely affect human health and safety. It is imperative to keep away these materials from humans [64]. Human operators' safety is also a critical issue in HRC, and a growing body of literature has investigated this context. [65] presents a comprehensive picture of different categories of safety in the HRC context. Moreover, [66] proposes an automated safety configuration for HRC regarding resources, processes, products, hazards, and the introduced safety behaviors model.

The distance between humans and cobots during operations is a significant factor that impact human safety. Generally, as the distance between a human and a cobot diminishes, the cobot gradually slows down until it eventually comes to a stop. Several researchers addressed the safe distance between humans and cobots. [67] develops simulation software that finds a safe distance for a human operator and a cobot by GA. Two criteria regarding human safety are integrated into the HRC disassembly planning model proposed in [1]. These criteria are preventing unsafe task allocation to the human operator and ensuring a safe distance between human and cobot operators. [68] balances an HRC assembly line by solving constraint integer programs (SCIP), focusing on maintaining a safe distance between a human and a cobot. This distance depends on several factors, such as human and cobot reaction times and the required distance for the cobot to stop. The proposed approach aims to allocate tasks in a manner, ensuring the distance between the human and the cobot is greater than the safe distance, leading to faster cobot operating and shorter process time.

2.5 Ergonomic risk in disassembly processes

Ergonomic risks in the workplace are a serious concern within the manufacturing industry. Failure to address these risks may result in work-related musculoskeletal disorders (WMSD) and other irreparable damage to the human skeletal system over time. According to the Occupational Safety and Health Administration (OSHA)[1], several activities at the workplace

[1]https://www.osha.gov/ergonomics/identify-problems

may potentially pose ergonomic risks. Applying overly strong force, frequently or continuously doing repetitive actions, and working in an inappropriate position are examples of these activities. Various approaches have been proposed to integrate ergonomic considerations into assembly/disassembly planning. Furthermore, in recent years and with the increasing use of cobots in manufacturing, a growing body of research has emerged to analyze the ergonomic risks in HRC industrial processes. [69] develops an HRC assembly line balancing model aiming to optimize both operation time and ergonomic risks. An integer planning model is developed in [70] to optimize assembly lines with respect to ergonomic risks. [71] introduces a fuzzy-based assembly line-balancing model that involves ergonomic risks. The proposed model considers four types of risk: twisting the wrist, lifting, twisting the hip, and squatting. [72] conducts an experiment that compares several key performance indicators (KPIs), including ergonomic risk, in the collaboration of a human operator and a cobot with a scenario that the human operator solely works. Results show that the HRC strategy effectively reduces ergonomic risks and physical stress compared to manual disassembly. A recent study [73] plans the HRC process by a hybrid ant lion optimizer according to multiple objectives, including the ergonomic consideration of humans, parts' recycling revenue, disassembly complexity, and operation time. The authors used a video processing approach to assess the ergonomic aspect. This approach extracts human joints from video frames to calculate the angles between the joints. Then, the rapid entire body assessment (REBA) approach measures the ergonomic factor based on the angles. In addition, the study models the process with an AND/OR graph approach.

2.6 Technical feasibility in disassembly processes

The feasibility of an operation is a major issue in a disassembly process. Operators sometimes require advanced tools, demolition machines, or even cannot disassemble some parts due to lack of strength. These parts remain incompletely disassembled and are transferred outside the cell. Therefore, it is essential to consider task feasibility in disassembly planning. One of the major drawbacks in the literature is that most introduced planning methods do not incorporate disassembly feasibility, and it has only been discussed in a few research studies. In [74], the authors have computed the feasibility and direction of contact and non-contact disassembly based on extracted information from CAD models. They also generated a weighted graph showing the contact type and feasibility in accordance with six axes. Moreover, some researchers [75, 76] consider feasibility based on the physical precedence of a product's component. In order to find the optimum sequence, GA then considers feasibility in addition to cost and environmental impacts as the objectives.

[77] introduces a partial destructive disassembly planning approach to address the line balancing problem. This approach considers the feasibility of destructive operations and the resulting noise pollution during the planning process. It presents a mathematical model with different objectives, including minimizing the number of stations, noise pollution, and related costs. Subsequently, an enhanced gray wolf algorithm is developed to optimize the task sequence. The MGFEM method [78] evaluates the disassembly feasibility of a product in a partial destructive process. First, it computes the failure characteristics of the parts based on experts' knowledge and the type of failure. Then, a total entropy value is calculated based on three significant factors in disassembly feasibility. If this value is higher than a threshold, partial destructive disassembly is not feasible. Alternatively, the algorithm moves forward and defines eight indexes for the most critical features in disassembly efficiency: component recognition, tool, direction, time, force, accessibility, the accuracy of positioning, and component size. These indexes are then quantified by a fuzzy-based approach. By applying an approach based on entropy, the algorithm computes weights for the indexes. Subsequently, it calculates the fine-grained comprehensive evaluation value. Finally, the components are sorted according to this value that determines their priorities.

2.7 Design for disassembly

In general, product design significantly affects disassembly processes by means of various factors. Several efforts have been made to identify these factors and the disassemblability rate of a design. It helps manufacturers to redesign products in a way that leads to a higher disassembly rate. Such a disassembly-aware design process that facilitates part recovery is called DfD. One of the first investigations in this area was presented in [79], in which the authors proposed an overall efficiency metric to evaluate product disassemblability. This value is computed using the information sheet of the product's components. The level of difficulty for product disassembly is defined based on its accessibility, required force, position, execution time, and other special metrics. Disassembly tasks are categorized into ten groups: pull/push, remove, unscrew, flip, cut, grip, deform, peel, pry out, and drill. The authors used a hair dryer as a case study for validating the proposed approach. After completing the corresponding sheet, they redesigned the product through four changes to improve disassemblability.

To compute the mentioned factors, other studies like [80] used a similar sheet-based approach to incorporate different characteristics of components in the estimated disassembly time calculation. The authors employed the Maynard operation sequence technique (MOST) proposed in [81] to determine the difficulty level for each component. They applied the proposed method to an electric drill as a case study. In a more recent study [82], Vanegas et al.

proposed ease of Disassembly Metrics (eDiM) to calculate disassembly time using MOST. In this approach, the disassembly process of each component consists of six fundamental tasks: tool change, identifying connectors, manipulation, positioning, disconnection, and removing. While tool change addresses taking and preparing a tool for an operation, identifying connectors refers to the time that an operator takes to find a connector location, its type, and a proper tool for its disassembly. Manipulation refers to the time for adjusting the product to properly position it for disassembly. Moreover, the action of putting a tool on a connector for its disassembly is referred to as positioning. Disconnection is the time required to disconnect a part from the product. Finally, removing concerns the time that an operator needs to remove a disassembled part to store it. The disassembly time for each component is calculated by summing the values of the six tasks, which are predefined in a database. Next, the disassembly time for every product component is entered into a table. A human operator then sums the relevant data to compute the overall optimized disassembly time for the given product, considering the disassembly precedence relationships.

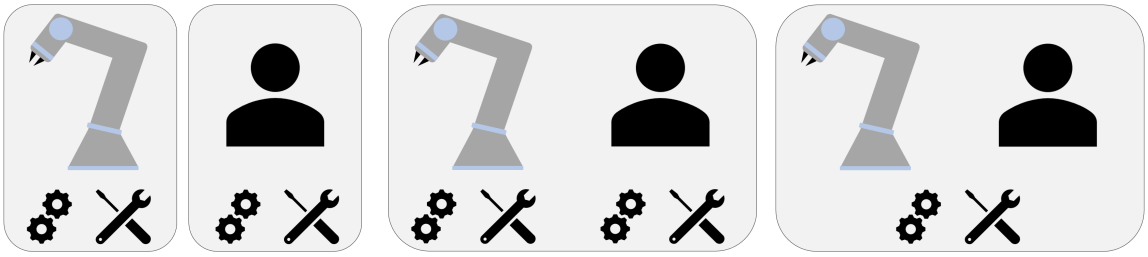
The main idea of LeanDfD introduced in [83] is to develop a disassembly sequence that considers time-based disassemblability and recyclability metrics in a recursive cascade process. Here, the time is calculated based on liaison types and properties through a liaison database. Moreover, another database containing material information is employed to calculate the recyclability metric. In the final stage, the computed time and recyclability metric are compared with two predefined threshold values. Upon satisfying both the time and recyclability thresholds, the proposed sequence is exported as a PDF/XML file. Otherwise, the process should be restarted.

2.8 Human-robot interactions

Over recent years, robots have been playing a growing role in industrial processes. Consequently, industries are undergoing significant evolution. Many manufacturing companies are becoming more and more interested in employing cobots that can work in parallel with human operators. The increasing evolution in the use of robots motivates AI researchers to investigate the field of robotics. According to a report on Nature's website [84], the number of research articles in the field of AI and robotics has grown significantly from 2015 to 2021.

In recent years, several efforts have been made to define human and robot interaction levels and methods. Rodríguez-Guerra et al. [85] have categorized all interactions between humans and robots into three main classes: coexistence, cooperation, and collaboration. As they explained, in coexistence scenarios, a human and a robot work individually on different tasks in separate workspaces. In contrast, a human and a robot work on distinct tasks but in the

same workspace in a cooperative environment. On the other hand, collaboration is defined as the condition that a human and a robot work simultaneously on the same task in the same workspace. All three scenarios are illustrated in Figure 2.2. As the authors pointed out, it is possible to add isolation and synchronization as two other classes to the main three ones.



(a) Co-existence (b) Co-operation (c) Collaboration

Figure 2.2 Human-robot interactions categories

In another research study [86], the authors categorized tasks performed by humans and robots in a collaboration cell based on an autonomy factor into four cases: leading, supportive, inactive, and an intuitive human and an adaptive robot. In the leading case, a human operator or a robot works autonomously on a task. In the supportive case, one of the actors provides assistance to the other during the operation. Moreover, a robot or a human operator waits for an upcoming task in the inactive mode. In the last case, they can switch their roles based on the situation at hand.

In this research, we address the collaboration scenario. In this case, as Figure 2.3 illustrates, a human operator and a cobot disassemble a product together in a cell. In this scenario, the capabilities of both actors are utilised to fulfill the tasks. Cobots can perform repetitive and unsafe tasks with high performance that may be hazardous for humans. However, experienced humans can react better to uncertainties in operations [1].

Various methods have been devised to address the HRC disassembly planning problem. Wurster et al. [87] presented a disassembly line-balancing model based on RL. The authors considered several manual, automatic, and autonomous workstations. Manual workstations are controlled by humans, while automatic workstations are controlled by robots. Autonomous stations have robots that can make their own decisions and learn from their mistakes. The authors utilised a Petri network and an RL architecture to model the entire pipeline and the decision-making process, respectively. They concatenated all information regarding product orders and workstation situations into a vector as the state of the RL model. In addition, the reward is proposed as a weighted linear combination of failed op-

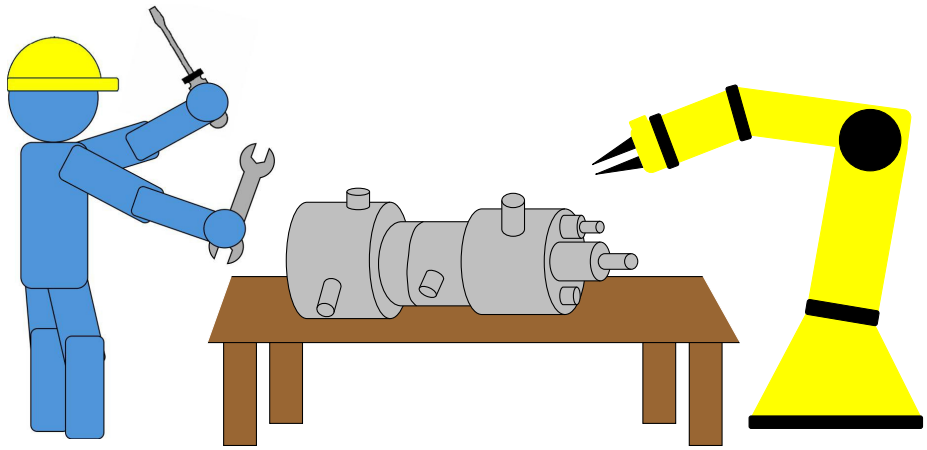


Figure 2.3 A human-robot collaboration cell

erations, disassembly time, and cost. A matrix-based HRC disassembly planning method has been proposed in [88]. The authors considered disassembly time as a predefined deterministic parameter. They designed a CAD model of an actual product in the first step. Then, they compute a precedence list between the product's components through a proposed matrix-based method. Afterward, multiple disassembly sequences are generated based on the precedence list of the product's components. Finally, the sequence with the shortest disassembly time is selected as the optimum disassembly sequence. In [1], the authors proposed an HRC disassembly planning approach that considers 14 criteria to formulate a sequential disassembly task allocation process into a numerical optimization-based problem. Human operators' safety is also considered in the decision-making process. The authors applied the proposed framework to a real case study (hard disk drive) to validate their approach.

Parsa and Saadat [86] have introduced an HRC disassembly planning model that represents the structure of an EoL product and the associated precedence relations by using an AND/OR graph. The proposed approach classifies tasks into eight different groups according to their corresponding difficulty and complexity. It optimizes the process by applying the GA, with the aim of allocating more challenging and heavy tasks to the human operator, while assigning easy and repetitive tasks to the cobot. The fitness function of the GA comprises several objectives, such as indices of untargeted components, operation time, and the frequency of human operator change. In [89], the authors proposed an HRC disassembly planning model based on a cascade structure. With the help of a digital twin framework, the model fuses data from real and virtual spaces. Then, it sequentially preprocesses the data to obtain parts features, types of liaison and tools, as well as disassembly precedence constraints. The model determines the most optimal sequence by using GA.

[90] introduces an HRC disassembly planning model based on a theoretical recursive learning algorithm. Initially, data from different sources is fused, extracting high-level information like human and cobot locations, hand gestures, body skeletons, and required tools. Subsequently, an RL model optimizes the process based on this information. Following that the operators execute the planned tasks. Finally, the acquired knowledge is shared using incremental learning and transfer learning techniques via a cloud-based technology. In the scope of lithium-ion battery recycling, [91] proposes a heuristic-based approach to assess the resilience of HRC disassembly based on several factors, such as stability, redundancy, efficiency, and adaptation. In the context of assembly planning, an RL-based HRC planning approach has been introduced in [92]. The proposed approach comprises a dual-agent model, an agent for a human operator and an agent for a cobot. An improved version of a deep deterministic policy gradient method is presented that computes a global Q-value based on the collaboration between two agents. The authors also represented the architecture of a product during the process by a simple vectorized approach. For evaluation, the proposed model was applied to the assembly of an alternator in a laboratory setting.

2.9 Joint applications of fuzzy logic and ML

Fuzzy logic is a classical data-driven approach that provides mathematical rules for modeling a diverse range of problems, particularly in scenarios with a high degree of uncertainty. On the other hand, ML refers to different techniques that efficiently and effectively predict patterns of data in numerous problems, including complex and large-scale scenarios. Hybrid fuzzy-ML models are becoming more and more popular in different applications as they combine the flexibility of fuzzy logic to tackle uncertain situations with the prediction power of ML models.

A considerable body of literature has investigated the incorporation of fuzzy-based methods with supervised and unsupervised models. [93] presents a fault diagnosis method for rotating machinery based on fuzzy clustering. It initially shifts raw vibration signals, captured from rotating machines, to the time-frequency domain, resulting in an image for each signal. It then utilises several image processing techniques, such as threshold filtering and neighborhood connectivity, to remove noise from images. Subsequently, it extracts latent features from images through Fourier descriptors followed by using fuzzy C-means to cluster the features. In another study [94], Sierra-Garcia and Santos have introduced a hybrid model based on deep learning and fuzzy logic for controlling wind turbine pitch. The proposed architecture contains two deep neural networks. While the first network estimates the current wind, the second network predicts future wind. The outputs of the two networks are subsequently

combined to represent effective wind. Finally, a fuzzy model sets the pitch control based on the effective wind and power error. In [95], the authors have developed a fuzzy-based system to control neural network parameters during the training phase. In a more recent paper [96], Walek and Fajmon have incorporated a fuzzy expert system into the architecture of a recommendation engine. The fuzzy system defines an importance level for each filtered item for user recommendations. In this regard, the fuzzy system assigns a level, either low, medium, or high, along with a descriptor, such as very, roughly, extremely, more or less, very roughly, very very roughly, rather, quite roughly, or significantly. In addition, several researchers [97–99] have incorporated fuzzy methods into neural networks' architectures to improve their interpretability.

Numerous research studies have investigated the combination of fuzzy logic and RL. One of the earliest works in this field, [100], proposes a hybrid RL and fuzzy neural network controller architecture that converges faster than conventional networks. In another earlier study [101], the authors developed a fuzzy Q-learning algorithm that outperforms the GA in the Cart-Centering problem. In a more recent study [102], Goharimanesh et al. have proposed a fuzzy RL (FRL) model to track trajectory in continuum robots. They tuned the parameters of the FRL model by the Taguchi method and the GA, resulting in an improved convergence time. [103] presents a hybrid learning and fuzzy logic model for autonomous vehicle planning. A Convolutional neural network (CNN) model and long short-term memory (LSTM) model combine sequential images of the surrounding environment and vehicle location information. Subsequently, the resulting vector is fed to a DQN as the state. The DQN outputs the steering angle and the accelerator. The authors considered the last layer's DQN values as membership degrees. They then defuzzified these values to generate continuous outputs for the steering angle and the accelerator. In recent years, several research studies [104–107] have explored the use of fuzzy reward functions in various environments. Using fuzzy values instead of crisp ones in the reward function enables an RL model to handle ambiguous and uncertain parameters in the environment's feedback more effectively.

CHAPTER 3 METHODOLOGY

In Chapters 1 and 2, we introduce the problem and review the corresponding literature. This chapter outlines the articles included in this thesis, presented in Chapters 4 through 7, to provide readers with a comprehensive perspective. Figure 3.1 illustrates the articles and their interrelationships. Following a comprehensive literature review, The first article proposes a real-time disassembly planning model based on multi-agent RL. Next, the second article presents a novel sustainable disassembly planning approach incorporating RL and fuzzy logic. This approach involves factors corresponding to all three sustainability pillars in decision-making. The third article assesses ergonomic risks associated with disassembly processes using video processing and fuzzy logic. Finally, the fourth article introduces a novel fuzzy-guided RL model for disassembly planning. The following sections provide a brief overview of these articles.

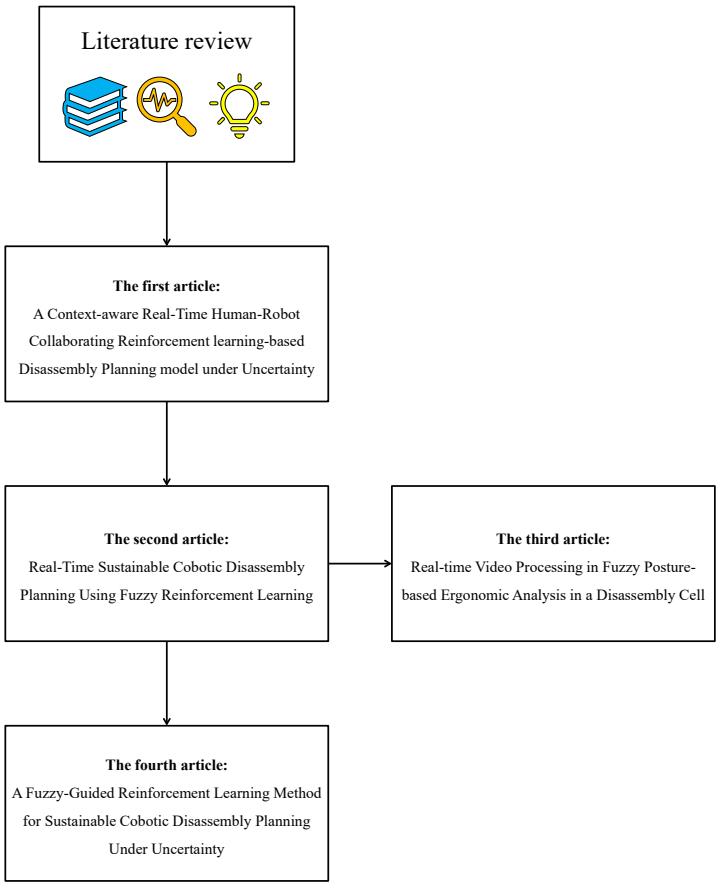


Figure 3.1 The presented articles and their interconnections

3.1 First article

As previously explained, a disassembly process contains uncertainties that may deviate it from the ideal path. Hence, it is crucial to plan the process dynamically with respect to online conditions. In this thesis, the first article develops an HRC disassembly planning model with real-time decision-making capability. Figure 3.2 illustrates an example of the proposed real-time task allocation approach. At t = 0, the model assigns tasks 1 and 2 to the human operator and the cobot, respectively. Task 3 is allocated to the human operator when task 1 is fully performed at t = T_1. Similarly, the model recommends task 4 to the human operator after task 3 is completed at t = T_1 + T_3. Following this, the model assigns task 5 to the cobot when task 2 is completed at t = T_2. This process will continue until operators complete all tasks. In this manner, the model assigns tasks according to the online circumstances rather than using a theoretically predetermined task sequence.

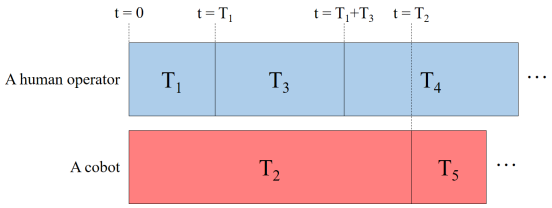


Figure 3.2 An example of the real-time task allocation mechanism

We model an EoL product architecture by using a graph-based approach illustrated in Figure 3.3. In this regard, each node and edge represent a component of the product and a connection between two components, respectively. We also define an origin node and assign a level to each node based on its distance from the origin. Figure 3.3 illustrates dotted line circles around the origin node, node 1. Nodes within an area between two circles have the same level. The disassembly process begins by selecting nodes from the outermost level. At each step, nodes from the current level, along with nodes connected to previously disassembled ones, are eligible for disassembly. The process progresses to the next level only when all nodes at the current level are disassembled. Notably, this graph-based approach applies a filter before node selection and does not directly select a node. In order to address this, we model the HRC disassembly problem in RL, enabling a learning-driven solution. We define the state as the product's online condition. Additionally, the action refers to the next product's component for disassembly. We also determine the reward function based on the problem's objectives, such as minimizing operation time and maximizing recovered parts quality.

Figure 3.4 illustrates the proposed multi-agent RL-based model that includes two agents— a

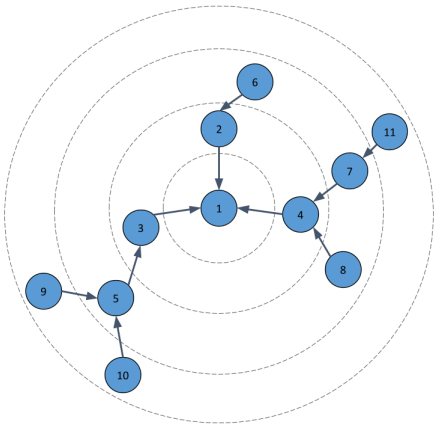


Figure 3.3 A graph of an EoL product

human and a robot— each selects tasks for its respective operator. Through dynamic interactions with an environment, these agents learn optimal policies for choosing the most efficient tasks to earn maximum cumulated reward values. It is a dynamic and recursive learning process, in which the agents adjust themselves based on feedback from the environment.

In the case of the human agent, after performing the action a_t^h in the state s_t^h, the agent receives the reward $r_{t'}^h$. Then, it moves from the current state s_t^h to a new state $s_{t'}^h$. The values s_t^h, a_t^h, $r_{t'}^h$, and $s_{t'}^h$ are stored in the replay memory M_h. Next, we sample a batch of data from the replay memory M_h to train the human agent. In this recursive process, the human agent is updated at each step. We consider the mentioned steps for the cobot too, in which all values s_t^r, a_t^r, $r_{t''}^r$, and $s_{t''}^r$ are stored in the replay memory M_r. This algorithm delivers the learned human agent and robot agent models that can make decisions in real-time.

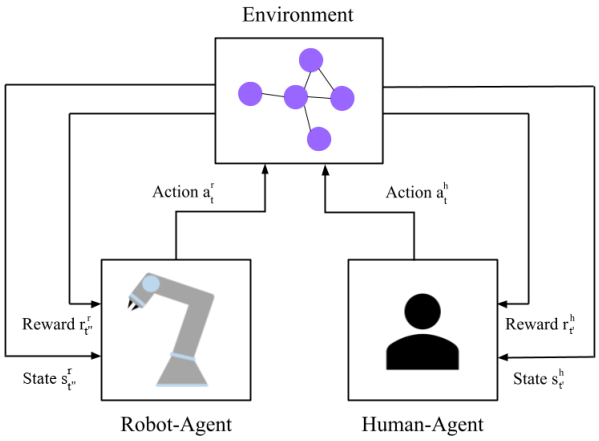


Figure 3.4 The proposed multi-agent RL-based model

3.2 Second article

In line with the move towards sustainable manufacturing, this chapter presents an RL-based disassembly planning model, derived from the definition of sustainability provided in Section 1.2, as shown in Figure 3.5. In this way, the model incorporates not only economic aspects but also social and environmental factors as the objectives in the optimization problem. Economic objectives include minimizing operation time and tool change frequency. In addition, minimizing ergonomic risk and maximizing human safety are social objectives. Meanwhile, the environmental objectives are to maximize the quality of recovered parts and minimize the cobot's energy consumption. The model also accounts for the technical feasibility of tasks in the planning process.

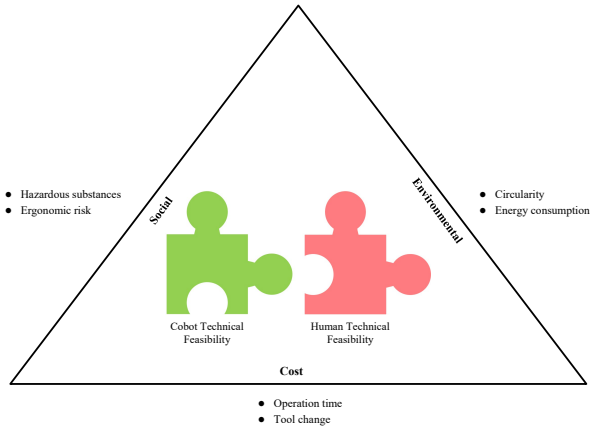


Figure 3.5 The proposed multi-agent RL-based framework

A disassembly process is associated with uncertainties that may appear in both the decision-making process and data. In terms of data, the mentioned factors result in the existence of uncertainty in parameters, such as the quality of recovered parts. Modeling these parameters is essential as they should be considered in the decision-making process. These parameters express subjective concepts, and consequently, it is difficult to numerically represent them due to the lack of prior knowledge about them. It is also impractical to determine exact values for these parameters because of their uncertain nature.

In order to represent data uncertainty and better align the model's training environment with real-world conditions, we developed a fuzzy-based environment that computes reward values in response to an executed action. It is notable that the reward functions consists of fuzzy and crisp terms. Difficulty, feasibility, operation time, and recovered quality are fuzzy parameters, while consumed energy, tool change, safety, and ergonomic risks are crisp parameters. We compute the operation time and recovered quality crisp values with a fuzzy-inference system

(FIS) according to difficulty and feasibility values. Then, we sum all of the crisp values to calculate the reward scores. As shown in Figure 3.6, the fuzzy-based environment includes three steps. It first executes the action selected by an RL agent. Subsequently, the FIS fuzzifies the uncertain parameters consequencing to the executed action. The FIS infers the data by using pre-defined fuzzy rules to compute outputs, which are then defuzzified. After that, the defuzzified parameters are combined with the non-fuzzy parameters based on the reward functions. Finally, the RL model receives the reward value from the environment.

Fuzzy numbers are commonly configured using expert opinions or historical data. However, in the second and fourth articles, due to the lack of access to such resources, the fuzzy numbers were logically set up to prove the functionality of the developed models.

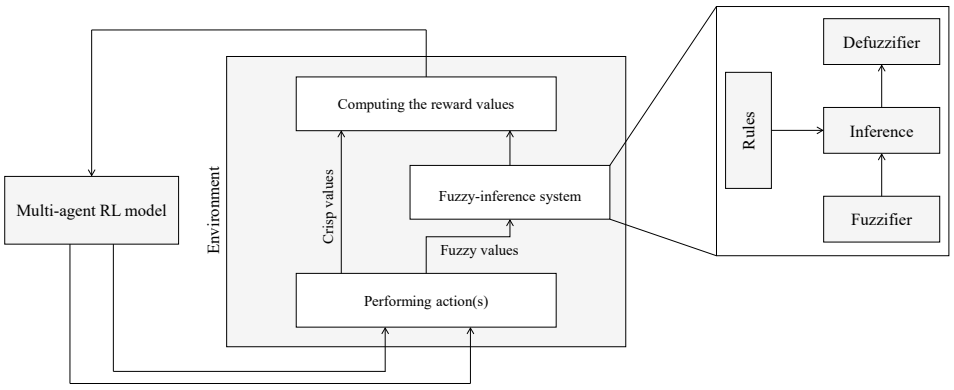


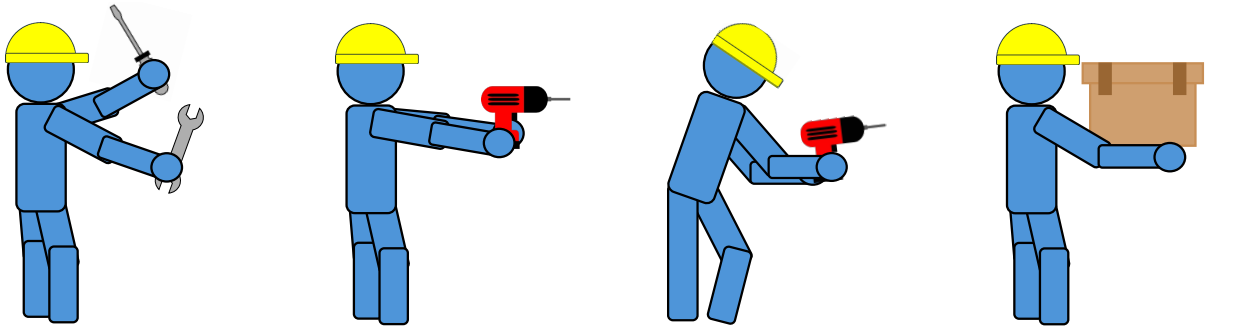
Figure 3.6 The fuzzy-based environment

3.3 Third article

Achieving a sustainable disassembly process requires considering human factors such as ergonomic risk and human safety. As described in Chapter 2, disassembly processes are associated with considerable ergonomic risk. According to Figure 3.7, performing repetitive and simple tasks, standing or sitting for a long time, working in an unstable position, and lifting heavy objects are examples of such high-risk tasks.

This research presents a novel approach based on image processing and fuzzy logic to assess the ergonomic risks of disassembly operations. This approach leverages a multi-camera structure, shown in Figure 3.8, with three cameras installed in front, right, and left of a human operator. This approach first extracts the upper body joints of the human in each frame. These joints are the left and right shoulders, elbows, hips, and wrists. The proposed approach then computes the angles between these joints, including those between the right and left wrists, elbows, and shoulders, as well as between the right and left elbows, shoulders, and

hips. We use three cameras as all angles cannot be calculated using videos taken with a single camera. Then, we evaluate the posture risk of each upper body part based on these angles through two directions: side-view and top-view. Following that, we define posture risks in elbows and shoulders by using rapid entire body assessment (REBA) and rapid upper limb assessment (REULA) approaches. In order to compute the final risk of each task, we involve the force/load of each task based on maximum voluntary contraction (MVC) to the posture risk. Subsequently, we compute the cumulative risk on each part based on the task duration compared to the task cycle time. Finally, we assess the total ergonomic risk of each part by using five fuzzy rules.



(a) Repetitive movements

(b) Long periods of sitting or standing

(c) Working in unstable positions

(d) Lifting heavy objects

Figure 3.7 Ergonomic risks corresponding to disassembly activities

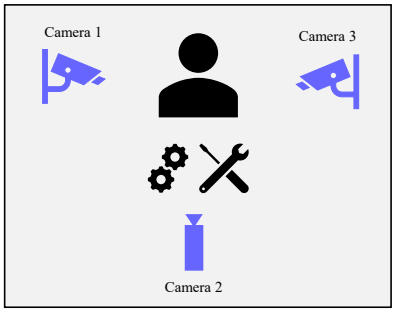


Figure 3.8 The proposed approach for ergonomic risk assessment

3.4 Fourth article

This chapter introduces an HRC disassembly planning model based on a multi-agent fuzzy-RL framework, illustrated in Figure 3.9. Each agent includes an RL model, a fuzzy model,

and an intelligence collaborative module. In each time step, an intelligence collaborative module selects a model to make decision. In this framework, RL models recursively learn with feedback from the environment. Meanwhile, fuzzy models are configured logically. Since RL models typically perform poorly and exhibit instability in early time steps, the fuzzy models are chosen more in initial iterations. In this case, the RL models learn with the feedback provided by the environment in response to decisions made by the fuzzy models. As the learning process evolves, the role of fuzzy models gradually diminishes, eventually allowing the RL models to make decisions independently. In this way, the fuzzy models play the role of copilot for the RL models. As the fuzzy model is developed with experts' knowledge, this study incorporates human expertise within an RL-based structure.

Furthermore, we conduct multiple sensitivity analyses to evaluate the model's performance under different uncertain conditions and scenarios with varying objective importance. These uncertain conditions are variable execution times for each task and task failure probability. In order to provide users with a convenient application, we have developed a GUI, shown in Figure 3.10. Using this interface, manufacturers can easily customize the process by adjusting each sustainable objective's importance according to their requirements. As shown in Figure 3.10a, by entering a number between 0 and 100 for each sustainability pillar (economic, social, and environmental), manufacturers can indicate the importance of the corresponding objectives in the decision-making process. Next, a pie chart illustrates the relative importance of the sustainability pillars, and task sequences for operators are generated, as shown in Figure 3.10b.

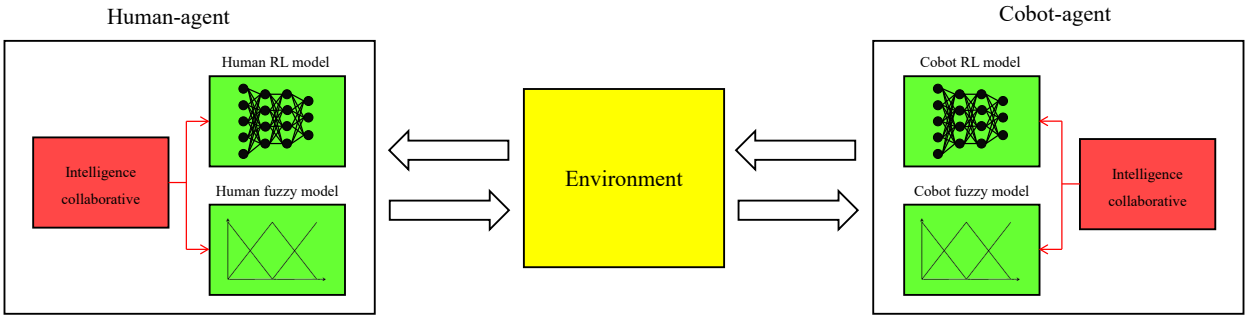
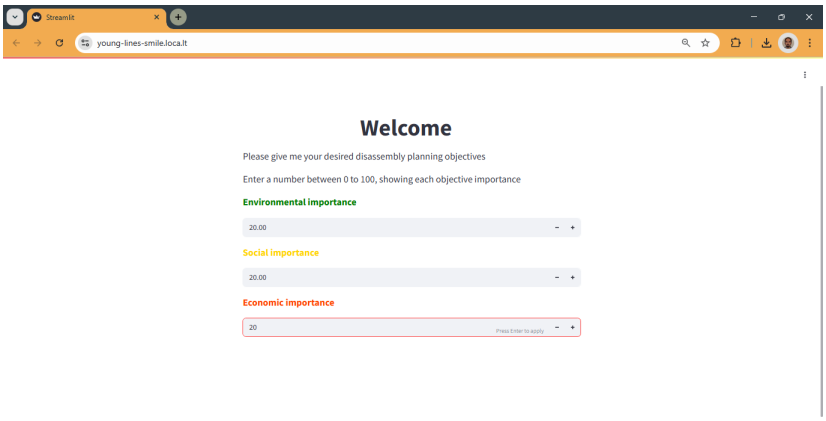
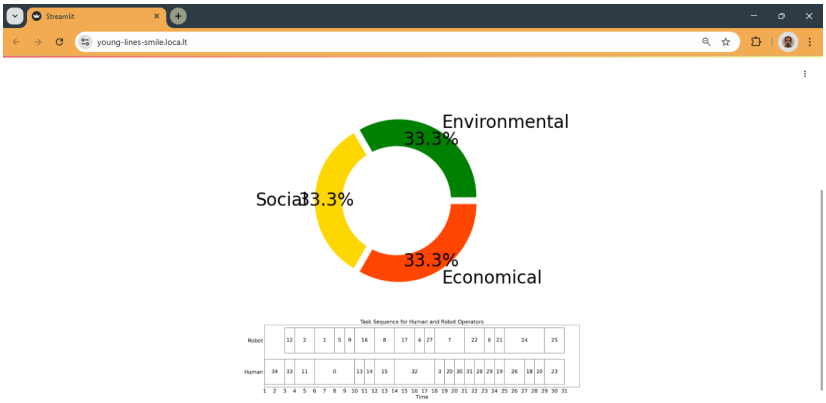


Figure 3.9 The multi-agent fuzzy-RL framework



(a) User inputs



(b) Customized outputs

Figure 3.10 The developed GUI

CHAPTER 4 ARTICLE 1: A CONTEXT-AWARE REAL-TIME HUMAN-ROBOT COLLABORATING REINFORCEMENT LEARNING-BASED DISASSEMBLY PLANNING MODEL UNDER UNCERTAINTY

Authors: Ashkan Amirnia and Samira Keivanpour

Submitted on March 4, 2023, and published on August 30, 2023, in *International Journal of Production Research (IJPR)*, Volume 62, Issue 11, pages 3972-3993, DOI: 10.1080/00207543.2023.2252526.

4.1 Abstract

Herein, we present a real-time multi-agent deep reinforcement learning model as a disassembly planning framework for human–robot collaboration. This disassembly plan optimizes sequences to minimize operation time and the disassembling costs of end-of-life (EoL) products. Combining different data-driven decision-making tools, the plan aims to handle the complexities and uncertainties of disassembly tasks. Based on the physical features and geometric limitations of EoL product components, we calculate product disassembly difficulty scores. Subsequently, the deep reinforcement learning model integrates these scores into planning process. The model allocates tasks in real time according to the online conditions of the human operator, cobot, and product, enabling the model to cope with uncertainties that may change the process routine. We also present different scenarios wherein a cobot collaborates with human operators with different skill levels. To evaluate model performance, we compare it with baseline models in terms of the convergence time and incorporated disassembly features. The analysis indicates that our model converges three times faster than a baseline model applied to the same case study. Moreover, our model includes more features of the disassembly problem in its decision-making process than any other baseline model.

Keywords: Disassembly planning; human-robot collaboration; real-time task allocation; reinforcement learning; context-aware recommender system

4.2 Introduction

Nowadays, many developed and developing countries are promoting recycling and waste reduction cultures due to economic and environmental benefits. As a result, there are an increasing number of disassembly methods that provide recovered product components for

reuse in manufacturing processes.

The use of robots in disassembly operations has grown in recent years. Collaborative robots (cobots) can enhance the quality of the disassembly processes in terms of completion time, human safety, and labor costs. Collaboration between a robot and a human operator reduces the disassembly process time, which leads to an increase in the production rate of manufacturing companies. Human operators are also exposed to ergonomic risks while performing disassembly tasks. In these cases, using cobots prevents jeopardizing human health and safety tremendously. Additionally, manufacturers may hire fewer human operators by using cobots during disassembly operations. Particularly, cobots can perform repetitive and simple tasks accurately, whereas a human operator may not accomplish them efficiently due to fatigue and distraction. However, task allocation in human-robot collaboration in disassembly processes has been challenging, and various methods have been developed to address this issue in the literature.

Zhao et al. [48] developed a reinforcement learning-based algorithm as a disassembly planning model. The authors introduced a new graph-based representation approach to obtain the precedences between product's components and consider the uncertainty of an EoL product. The authors have presented an algorithm to assign a level to each component of a product based on their positioning within the hierarchical structure. Accordingly, their task allocation process begins with the components at the outermost level. After disassembling all the components at this level, it moves to the next inner level. In this way, the process continues until all the components have been disassembled and the innermost level is reached. The authors also proposed a reward function. It takes into account processing time, profit, and the difference in the levels of the previous and newly selected components in a disassembly sequence. The last term awards the agent once it completes the disassembly of components in a level and moves to the next level.

In another paper [88], a matrix-based human-robot collaboration disassembly planning method has been proposed. The authors considered disassembly time as a predefined deterministic parameter. They designed a CAD model of an actual product in the first step. Then, they compute a precedence list between the product's components through a proposed matrix-based method. Afterward, multiple disassembly sequences are generated based on the precedence list of the product's components. Finally, the sequence with the shortest disassembly time is selected as the optimum disassembly sequence.

Lee et al. [1] proposed a human-robot collaboration disassembly framework. It considers 14 criteria to formulate a sequential disassembly task allocation process into a numerical optimization-based problem. Human operators' safety is also considered in the decision-

making process. The authors applied the proposed framework to a real case study (hard disk drive) to validate their model.

Parsa and Saadat [86] have introduced cleanability, repairability, and economy metrics as three disassemblability factors. These metrics are defined by human experts and are implemented to prioritize components in partial human-robot collaboration disassembly planning. In addition, they have assigned various weights to the metrics to determine their importance in the prioritizing process. All disassembly operations are categorized into eight groups with different numerical score based on their difficulty and complexity. The disassembly framework's objective is to allocate difficult and complex tasks to a human operator that a robot cannot perform due to its physical limitations. Moreover, an AND/OR graph-based approach represents the entire disassembly process and precedence hierarchies. Then, a genetic algorithm is employed to optimize the task allocation process, which has a fitness function based on three variables: disassembly time, non-targeted component index, and human operator change. The last two variables play the role of penalty terms in the fitness function. They have applied their proposed method to a fuel pump as a case study for evaluation.

In the context of assembly planning, a reinforcement learning-based human-robot interaction model has been introduced in [92]. The authors developed a multi-agent framework that considers an agent for a human operator and an agent for a robot. An improved version of a deep deterministic policy gradient approach is presented as well. It computes a global Q-value based on the collaboration between two agents. The authors also utilised a simple vectorized approach to represent the architecture of an assembly product. For evaluation, the proposed model was applied to the assembly of an alternator in a laboratory setting. Similarly, a reinforcement learning-based assembly planning model has been introduced in [49]. The authors considered two vectors. First, a binary vector represents whether the assembly task is finished or not. Secondly, a vector containing the required assembly tool as the state. The reward signal consists of two parts: completion time and user satisfaction. The latter is defined based on how satisfied the user is with the quality of assembly sequences. Further, the disassembly completion times for each component can be configured both deterministically and stochastically. In the proposed framework, the authors compared several reinforcement learning models that were applied to airplane toy data provided in [50, 51].

Recent research utilised static modeling approaches that cannot operate dynamically to adapt to the disassembly process's uncertainties and perform the decision-making process in real time. As [2] points out, the optimal disassembly sequence should be updated based on the actual state of the disassembly process through the development of revised techniques. Moreover, as discussed in [48], it is generally impossible to utilise fixed disassembly sequences for

end-of-life (EOL) products due to their uncertain structure. Furthermore, if minor structural alterations are made to EoL products, heuristic approaches may need to reevaluate their optimization process. Thus, structural variations in EoL products cannot be addressed by heuristic methods. Similarly, according to [2], proposed static approaches in most of the previous research cannot adapt dynamically to uncertainties in real disassembly scenarios.

In addition, most of the previous research did not consider the disassembly difficulty level for each part or component in the task allocation procedure. Moreover, previous studies have not addressed the skill levels of human operators, which may influence the entire process performance. This paper aims to fill these gaps by proposing a context-aware recommender system for optimizing disassembly sequence planning with a cobot.

The study's main contributions are outlined as follows:

- Propose a new disassembly planning model for the human-robot collaboration process. The model performs the decision-making process in real time rather than generating a predetermined series of tasks. This feature allows the model to cope with the process uncertainties well.
- Develop a multi-agent model containing two agents: the human agent and the robot agent. In this case, each agent allocates tasks to its corresponding operator.
- Propose a new state for the process according to the disassembly difficulty scores. We use several physical and geometric features of a product's components to compute these scores.
- Represents the architecture of an EoL product by a new graph-based approach. We assign a level to each component of the product based on its positioning in the graph.
- Propose an algorithm for switching from the multi-agent model to the single-agent model, in which the human operator performs the tasks solely. This approach is employed whenever there is no appropriate task for the cobot, and performing a task can break its physical structure.
- Involve skill levels of human operators in the task allocation process. It enables the model to recommend relevant tasks to human operators.

The rest of this paper is organized as follows. Section 4.3 explains the context of the problem and the relevant research fields. Next, Section 4.4 describes the proposed model. Section 4.5 includes all experimental setups, results, and discussion. Finally, Section 4.6 summarizes the paper and outlines the future steps for this research study.

4.3 Background

4.3.1 Disassembly processes

As described in [2] and [108], remanufacturing involves utilizing sustainable reuse methods rather than recycling, landfilling, or incineration of EoL products. The first step toward remanufacturing is to take apart EoL products to extract their components. Whenever possible, these components are repaired and used in the manufacture of new products.

Zhou et al. [2] categorized the objectives of disassembly processes into four groups: cost, profit, ecological effects, and practical purposes. According to the authors, the cost function includes process duration, process equipment cost, labor cost, number of parts, number of operations, maintenance degree provided in [109], mean time to repair (MTTR) index introduced in [110], and the traveled distance to disassemble a component. The authors also defined the profit for an EoL product as a function of the components' profit, retrieved weight or volume, profit probability, variable product value, and retrieved value indexes. The ecological effects of a part and the impact of a disassembly procedure on the environment are also considered as two environmental indexes by authors.

There is a growing body of literature exploring how to optimize a disassembly process by sequential task allocation. [33] proposes a hybrid optimization method as a disassembly planning model that combines the Genetic algorithm with the Tabu search. It also utilises a graph-based method called disassembly constraint graph (DCG) to represent an entire product's structure. In another study, [34] introduces an approach based on the Genetic algorithm for disassembly line balancing. Next, a disassembly sequence planning approach based on the ant colony optimization algorithm is presented in [31]. The authors formulated the problem using three factors: the number of components, the disassembly tool, and the disassembly process direction. [16] tackles the disassembly sequence planning problem by defining two component classes. The first class applies Dijkstra's algorithm only to functional components, while the second one runs a PSO-based algorithm for all components, including fasteners to obtain the most optimal disassembly arrangement.

Some preliminary work was carried out in this field utilizing machine learning approaches to plan a disassembly process. [44] proposes a dynamic programming approach for disassembly planning that handles the uncertainty of the process. The author has employed a Petri network to model the entire process. In addition, a combination of a Petri network and a hybrid Bayesian network is proposed as a disassembly planning model in [40]. The Petri network is utilised for representing the entire process, and the hybrid Bayesian network selects the next action in the process.

In general, product design significantly affects disassembly processes by means of various factors. Several efforts have been made to identify these factors and the disassemblability rate of a design. It helps manufacturers to redesign products in a way that leads to a higher disassembly rate. Such a disassembly-aware design process that facilitates part recovery is called design for disassembly (DfD). One of the first investigations in this area was published in [79]. They have presented an overall efficiency metric to evaluate product disassemblability. This value is computed using the information sheet on the product's components. The level of difficulty for product disassembly is defined based on its accessibility, position, required force, required time, and other special metrics. Disassembly tasks are categorized into ten groups: pull/push, remove, unscrew, flip, cut, grip, deform, peel, pry out, and drill. They considered a hair dryer as a case study for the proposed approach. After completing the corresponding sheet, they redesigned the product through four changes to improve the disassemblability feature.

To compute the mentioned factors, other studies like [80] used a similar sheet-based approach to incorporate different characteristics of components in the estimated disassembly time calculation. The authors employed the Maynard operation sequence technique (MOST) proposed in [81] to determine the difficulty level for each component. They applied the proposed method to an electric drill as a case study. In a more recent study, [82] proposes ease of Disassembly Metrics (eDiM) to calculate disassembly time using MOST. In this approach, the disassembly process of each component consists of six fundamental tasks: tool change, identifying connectors, manipulation, positioning, disconnection, and removing. While tool change addresses taking and preparing a tool for an operation, identifying connectors refers to the time that an operator takes to find a connector location, its type, and a proper tool for its disassembly. Manipulation refers to the time for manipulating the product to properly position it for disassembly. Moreover, the action of putting a tool on a connector for its disassembly is referred to as positioning. Disconnection is the time required to disconnect a part from the product. Finally, removing concerns the time that an operator needs to remove a disassembled part to store it. The disassembly time for each component is obtained by summing up the associated value for each of the six tasks in a predefined database. Next, the disassembly time for every single component of a product is imported into a table. Then, a human operator adds up the relevant imported data to get the overall optimized disassembly time for a given product, considering the disassembly precedence relationships.

The main idea of LeanDfD introduced in [83] is to develop a disassembly sequence that considers time-based disassemblability and recyclability metrics in a recursive cascade process. Here the time is calculated based on liaison types and properties through a liaison database. Also, a database containing material information is employed to calculate the recyclability

metric. In the final stage, the computed time and recyclability metric are compared with two threshold values. Upon satisfying both the time and recyclability thresholds, the proposed sequence is exported as a PDF/XML file. Otherwise, the process should be restarted.

On the other hand, researchers have become interested in computer vision and image processing systems incorporated into disassembly processes. [53] proposes an image processing-based algorithm for screw detection in electric vehicle motors. This algorithm integrates several image processing techniques, including HSV color modeling, image depth detection, grayscale conversion, and the Harris detector proposed in [54]. In another research study, [55] introduces an approach based on Tiny-Yolo v2 to detect screws for robot operators in disassembly processes.

Furthermore, there has been growing attention to utilizing reinforcement learning models for assembly and disassembly task scheduling problems over the past few years. These models operate dynamically and adaptively considering any change in products' structures. The state for these reinforcement learning models is defined as the product's situation at each step. Moreover, the action is assumed to be the assembly/disassembly of the next component. An early investigation of reinforcement learning for the purpose of disassembly planning was performed in 2010 in [43]. The authors presented a reinforcement learning-based model for disassembly line planning that utilises an Eleman network for Q-value approximation. Mao et al. [45] also introduced a hybrid DQN model that utilises the genetic algorithm to improve the long-term reward for disassembly sequence planning. A Petri network is also used to represent the product's architecture. The authors evaluated the proposed model for maintenance training in a virtual reality environment.

4.3.2 Human-robot interactions in disassembly planning

Over recent years, robots have been playing a growing role in industrial processes. Consequently, industries are undergoing significant evolution. Many manufacturing companies are becoming more and more interested in employing collaborative robots (cobots) that can work in parallel with human operators. The increasing evolution in the use of robots motivates AI researchers to investigate the field of robotics. According to a report on Nature's website [84], the number of research articles in the field of AI and robotics has grown significantly from 2015 to 2021.

In recent years, several efforts have been made to define human and robot interaction levels and methods. [85] categorizes all interactions between humans and robots into three main classes: coexistence, cooperation, and collaboration. As they explained, in coexistence scenarios, a human and a robot work individually on different tasks in separate workspaces.

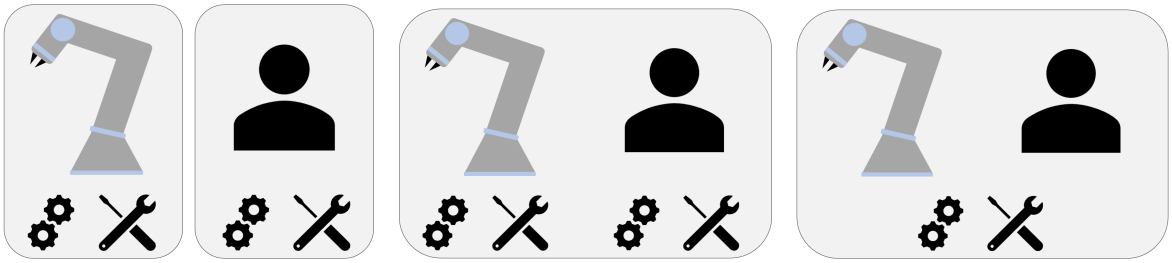
In contrast, a human and a robot work on distinct tasks but in the same workspace in a cooperative environment. On the other hand, collaboration is defined as the condition that a human and a robot work simultaneously on the same task in the same workspace. All three scenarios are illustrated in Figure 4.1. Furthermore, three main categories can be extended by including isolation, synchronization, or other classes.

In another research study, Parsa and Saadat [86] categorized tasks performed by humans and robots in a collaboration cell based on an autonomy factor into four cases: leading, supportive, inactive, and an intuitive human and an adaptive robot. In the leading case, a human operator or a robot works autonomously on a task. In the supportive case, one of the actors provides assistance to the other during the operation. Also, a robot or a human operator waits for a upcoming task in the inactive state. In the last case, they can switch their roles based on the situation at hand.

Additionally, manufacturing operations such as disassembly processes are changed through collaboration between humans and robots. In this scenario, the capabilities of both actors are utilised to fulfill the tasks. Robots can perform repetitive and unsafe tasks with high performance that may be hazardous for humans to operate. However, experienced humans can react better to uncertainties in operations as explained in [1].

Various methods have been devised to address the human-robot collaboration disassembly planning problem. [87] presents a disassembly line-balancing reinforcement learning-based model. The authors considered several manual, automatic, and autonomous workstations with products divided between them. Manual workstations are controlled by humans, while automatic workstations are controlled by robots.

Autonomous stations have robots that can make their own decisions and learn from their mistakes. The authors utilised a Petri network and a reinforcement learning architecture to model the entire pipeline and the decision-making process. They concatenate all information



(a) Co-existence (b) Co-operation (c) Collaboration

Figure 4.1 Human-robot interaction categories

regarding product orders and workstation situations into a vector as the state. Finally, the reward is proposed as a weighted linear combination of failed operations, disassembly time, and cost.

Researchers are becoming increasingly interested in sustainable development in this field. As an example, [90] introduces a sustainable manufacturing framework. The researchers proposed a theoretical recursive process containing five modules. It includes multi-modal perception, multi-target cognition, strategy and decision-making, control and execution, and knowledge formation and evolution. The multi-modal perception module collects data from different sources. It includes physical, transport, cyber, and application layers. Multi-target cognition module imports data to detect disassembly tools, and human and robot positions. It will then be fed to the decision-making stage for further processing. A dynamic approach based on reinforcement learning is proposed for the strategy and decision-making module that will be trained on a digital twin platform. After running the control and execution module, incremental and transfer learning is used to pass knowledge among robots through clouds via the knowledge formation and evolution module.

4.3.3 Deep reinforcement learning in manufacturing

Integrating the prediction power of deep neural networks with the self-adaptive feature of reinforcement learning makes deep reinforcement learning models strong tools to solve decision-making problems under uncertainty. Optimizing the manufacturing process by deep reinforcement learning techniques is an emerging research field. Over the past few years, several researchers have addressed deep reinforcement learning applications in this domain. [111] proposes an intelligent scheduling and reconfiguration framework based on a multi-agent deep reinforcement learning model for a reconfigurable flow line. The multi-agent model includes the scheduling agent and the reconfiguration agent. While the scheduling agent selects a job for unoccupied machines, the reconfiguration agent defines which workshop settings should be reconfigured. In the context of cloud manufacturing, [106] presents a reinforcement learning model for task assignment in a multi-project scheduling problem. Next, [112] proposes a dynamic scheduling approach for robot services based on a deep reinforcement learning framework. In this regard, the authors used both DQN and Dueling DQN (DDQN) in the framework to provide a comparison analysis. In another research study, Sakr et al. [113] introduced a multi-agent reinforcement learning-based model to dispatch and allocate resources in a complex semiconductor manufacturing system. In order to maximize the performance of the aerospace supply chain, [114] proposes an inventory replenishment approach based on a hybrid model involving simulation and reinforcement learning. Next, [115] plans aero-

engine fleet maintenance by concatenating the Q-values of all engines' maintenance actions. Herein, the authors proposed a multi-agent reinforcement learning model in which each agent computes the corresponding Q-value vector of each engine.

In recent years, deep reinforcement learning methods have become increasingly popular in human-robot interaction. [116] introduces a multi-agent reinforcement learning model based on the DQN architecture to optimize the human-robot collaborative assembly problem. In this regard, each agent represents the role of a cobot or a human operator in the decision-making process. [117] addresses safe human-robot interaction by developing a reinforcement learning model for motion planning to avoid accidents. Liu et al. [118] have developed a multi-agent reinforcement learning-based framework for the human-robot collaborative assembly process. The main goal is to train a cobot to adaptively work with a human operator in an interactive assembly environment. In another research study, [119] proposes a model based on the Actor-Critic architecture for safe human-robot collaboration in industrial environments.

4.3.4 The required disassembly features comparison and the potential gaps

This part aims to compare several studies in terms of the incorporated disassembly features in order to discover potential gaps in the literature. Due to the lack of numerous disassembly planning models, we evaluate our approach against two assembly task allocation models as they have similar features compared to ours.

Table 4.1 shows the papers and their corresponding features. All the papers lack the real-time decision-making feature. As previously mentioned, a process's routine may be changed by unexpected incidents. Hence, it is significant for a planning model to perform the decision-making in real time instead of using a predetermined task sequence. Next, the proposed models in (4,5,7) utilise the difficulty factor for disassembly tasks as an influential element in the decision-making process. Although, the difficulty level factor is taken into account without addressing the products' physical characteristics (5,7). The difficulty is a critical factor in disassembly task allocation. Extremely difficult tasks cannot be efficiently executed by low-skilled human operators. Furthermore, cobots cannot perform difficult tasks because of their physical limitations. Therefore, inefficient task allocation results in longer processing time and lower recovered parts' quality.

Although all papers consider the time of the process as a noteworthy feature, they do not integrate human operators' skill levels and recovered part quality into their models. Assembly/disassembly operations are skilled-based jobs. A human operator's skill level can affect the output quality of an operation. Several papers in the field of manufacturing consider the skill levels of human operators as an assumption in modeling processes. In [120] and [121], the

authors integrate human operators' skill levels in an assembly line balancing process by considering three types of skilled workers: low-skilled, medium-skilled, and high-skilled. In [122], the authors involve proficiency levels and skill sets as influencing factors in a disassembly line balancing optimization model. Since the disassembly target is to use the recovered parts for new product manufacturing, the quality of these parts is also an important factor. Further, the models presented in (2,3,4,7) are capable of self-adaption, which is critical for adjusting to dynamic situations in assembly/disassembly cells. Finally, (2,3,4,5) proposed planning models for the collaboration of a human operator and a cobot in assembly/disassembly cells.

Even though the previous studies examined different features of disassembly processes, the proposed models can still be improved. Therefore, this article introduces a comprehensive disassembly planning model to cover different aspects of the problem. In this paper, we introduce a human-robot collaboration disassembly planning model that allocates tasks in real time. It is a self-adaptive model that aims to minimize operation time and maximize recovered part quality. We also incorporate several features related to difficulty and human operators' skill levels into the planning process.

4.4 Methodology

This paper proposes a novel interactive disassembly planning model for human-robot collaboration, where a human operator and a cobot work alongside each other in a shared work cell. This planning model aims to allocate tasks to operators depending on several physical and geometric features of EoL products, human operators' skills, and cobots' abilities.

The outline of this section is organized as follows: the architecture of an EoL product, its components, and the connections between them are discussed in Section 4.4.1. Next, Section 4.4.2 presents the concepts of the task allocation process. A disassembly process is modeled based on a reinforcement learning framework in Section 4.4.3. We discuss the switching between collaborative and inactive modes in Section 4.4.4. The real-time task allocation

Table 4.1 Comparison of required disassembly features in the various models

Index	Paper	Real-time	Difficulty	Operation time	Human skill levels	Quality	Self-adaptive	Cobot
1	[48]	×	×	✓	×	×	✓	×
2	[88]	×	×	✓	×	×	×	✓
3	[1]	×	×	✓	×	×	×	✓
4	[86]	×	✓	✓	×	×	×	✓
5	[92]	×	✓	✓	×	×	✓	✓
6	[49]	×	×	✓	×	×	✓	×
7	[45]	×	✓	✓	×	×	✓	×
8	Our model	✓	✓	✓	✓	✓	✓	✓

process is explained in Section 4.4.5. Finally, we explain our proposed multi-agent deep reinforcement learning algorithm for the task allocation process in Section 4.4.6.

4.4.1 A product architecture modeling

EoL products have diverse shapes and geometries. They are composed of several components. The first step in an EoL product disassembly is to model the entire product architecture to analyze the disassembly precedence relationships. This step provides a represented form of a product for the decision-making process. As discussed in [2], graphs, Petri networks, matrices, and universal methods are the four classes of disassembly representation models. As a contribution, we utilise a directed acyclic graph-based approach to represent a product architecture and the connections among its components. Each node of a graph refers to each component, and the graph's edges represent the connections between the components. We also employ the depth-first search algorithm to assign a numerical score as a level to each node based on its position in the product's architecture.

Figure 4.2 depicts a simple product graph. The illustrated dotted line circles define the nodes' levels, in which nodes within an area between two circles have the same level. These levels are arranged in ascending order. Accordingly, the innermost and outermost nodes have levels one and four, respectively.

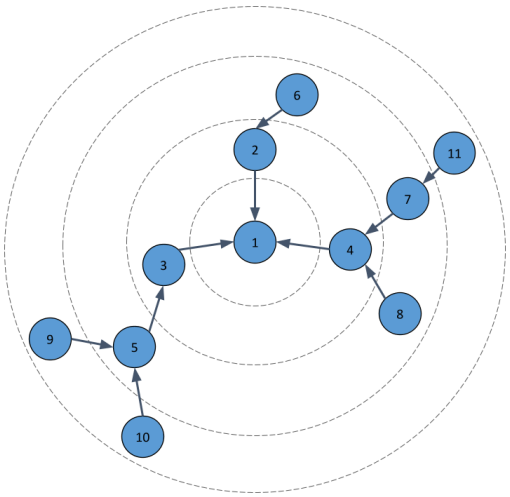


Figure 4.2 A graph of an EoL product

Based on these obtained levels, we provide the Level vector that includes all nodes' corresponding levels. This vector is shown in Equation 4.1, where M and l_i represent the number of the nodes and the level of node i, respectively. Moreover, Equation 4.2 denotes the Adjacency matrix of the graph. It is a square M $\times$ M matrix that each of its elements refers

to a connection between two nodes. In this regard, the element $a_{i,j}$ identifies the connection status between nodes i and j. As shown in Equation 4.3, the value of $a_{i,j}$ is one if a connection exists between nodes i and j. Otherwise, it is zero.

$$\text{Level vector} = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_M \end{pmatrix} \tag{4.1}$$

$$\text{Adjacency matrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,M} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \cdots & a_{M,M} \end{pmatrix} \tag{4.2}$$

$$a_{i,j} = \begin{cases} 1 & \text{if there is an edge between node } i \text{ and node } j \\ 0 & \text{otherwise} \end{cases} \tag{4.3}$$

The decision-making process to select a node should be performed according to the precedence relationships. For example, node 3 in Figure 4.2 cannot be disassembled at the beginning of the process without disassembling nodes 9 or 10, and 5. Therefore, we provide a fresh method to consider limitations imposed by the precedence relationships in the decision-making process through the Level vector and the Adjacency matrix.

Algorithm 1 presents the decision-making process concerning the precedence relationships. First, we select a node among the nodes at the highest level. During the while loop, we select a node at the highest level or connected to the first selected node. Upon selecting all nodes at the current level, the algorithm proceeds to the next level by reducing one level. Finally, the algorithm breaks the loop if the current level is zero or if all nodes of the product are disassembled.

Notably, this algorithm only applies the limitations of the precedence relationships to the decision-making process. We utilise a learning-based approach explained in the following sections to select nodes.

Algorithm 1 The decision-making process with respect to the limitations of the precedence relationships

Input: the Level vector, the Adjacency matrix

Select a node with the highest level.
while true **do**
  Select the nodes in the current level or nodes connected to the disassembled nodes.
  if all nodes in the current level are disassembled **then**
    level = level - 1.
    if level = 0 || all nodes are disassembled **then**
      Break the loop.
    end if
  end if
end while

4.4.2 Task planning strategy based on the disassembly difficulty

Manual disassembly tasks performed by human operators require a high level of skill. Inexperienced human operators may not have enough skills to employ the required tools during the processes effectively. It can lead to longer processing times and lower recovered part quality. Specifically, they cannot perform more difficult and challenging operations efficiently. More experienced humans, on the other hand, can usually complete tasks in less time and with higher quality. We consider human operators' work experience as an influential feature in the process's performance. Towards this end, we categorize human skill levels based on their work experience into three levels: acceptable, medium, and high.

On the other hand, cobots can do simple and repetitive tasks that a human operator might find distracting. In contrast, they cannot accomplish difficult tasks that require high strength and flexibility. Therefore, it is recommended to allocate these challenging tasks to human operators.

To recommend a related task to each operator, we categorize the tasks based on their disassembly difficulty and complexity into three levels: easy, difficult, and very difficult. In this regard, easy tasks are assigned to cobots. Human operators with acceptable-skill levels are preferred for performing easy tasks too. Also, human operators with medium-skill levels can handle easy and difficult tasks efficiently. Ultimately, highly experienced human operators can perform any type of task. All operator types and their corresponding tasks' difficulty levels are illustrated in Table 4.2.

Table 4.2 Operators' types and their proper tasks in terms of disassembly difficulty

Operator type	The difficulty levels of appropriate tasks
Cobot	Easy
Acceptable-skill human operators	Easy
Medium-skill human operators	Easy Difficult
High-skill human operators	Easy Difficult Very difficult

4.4.3 Modeling a collaborative disassembly problem based on reinforcement learning

We formulate the disassembly planning problem as a standard multi-agent reinforcement learning process. It includes an environment and two agents: the human agent and the robot agent. In this way, the human agent and the robot agent perform the decision-making processes to select tasks for a human operator and a cobot, respectively.

Each agent starts from an initial state s_t at step t. It then performs an action a_t. It moves to a new state s_{t+1} by receiving feedback from the environment as a numerical reward r_{t+1} at step t+1. As it is a recursive process, agents learn a policy π during these interactions with the environment to select an action in each state. Consequently, this process continues until agents learn the optimal policy, allowing them to select actions with the maximum accumulated rewards. Equation 4.4 illustrates the optimal policy, where the terms $\pi^\star$ and T address the optimal policy and the total time steps, respectively. The term r_t also refers to the reward at step t.

$$\pi^\star = \arg\max_{\pi} \mathrm{E}\left[\sum_{t=0}^{T} r_t\right] \tag{4.4}$$

To formulate the agents-environment interactions, we utilise the Markov decision process (MDP). It is a tuple containing four elements: s_t, a_t, r_{t+1}, and s_{t+1}. As [123] explains, MDPs can model detailed theoretical expressions since they are the mathematically optimal presentations of reinforcement learning problems.

State

This study aims to integrate the disassembly difficulty of components into the decision-making process. Therefore, we define the state as a difficulty vector, an array including the disassembly difficulty scores of all components.

To compute these disassembly difficulty scores, we use the physical and geometric features

of a product's components. In this regard, we consider the weight and size of a component, required strength for disassembling, destructive or non-destructive disassembly tasks, the number of its connections with other components, and the liaisons' types and properties as the influential physical features of a product's disassembly difficulty.

Furthermore, accessibility and positioning are two major geometric features. As explained in [80], accessibility refers to the ability for accessing to a part by hand or tool. Also, [82] defines positioning as placing a required tool on a component's connection before disassembling it.

Assigning numerical values to each of these physical and geometric features is an essential step for utilizing them. For this purpose, we utilise numerical values for strength, weight, size, positioning, and accessibility provided in [86]. We also determine the destructive feature as a binary variable. In this case, destructive and non-destructive disassembly tasks are labeled one and zero, respectively. Similar to [83], we assign a discrete value to each liaison type and a value between one and two to each liaison's properties.

Then, a linear combination of these seven features is defined to address the difficulty score of a component's disassembly. Equation 4.5 demonstrates the linear combination, where w_1, w_2, ... and w_7 are the corresponding features' coefficients. These coefficients indicate the importance of each feature and are defined by experts. Moreover, a product's disassembly difficulty vector D is illustrated in Equation 4.6, where d_i represents the disassembly difficulty score of component i. Notably, the defined disassembly difficulty vector is considered as the states of both agents.

$$\begin{aligned} d_i = {} & w_1 \cdot \text{strength} + w_2 \cdot \text{weight} + w_3 \cdot \text{size} + w_4 \cdot \text{destructive} \\ & + w_5 \cdot \text{positioning} + w_6 \cdot \text{accessibility} \\ & + w_7 \cdot (\text{liaisons scores} \cdot \text{liaisons properties} \cdot \text{number of connections}) \end{aligned} \tag{4.5}$$

$$D = [d_1, d_2, d_3, \ldots, d_i, \ldots, d_N] \tag{4.6}$$

Action

The model's purpose is to allocate the most suitable components for disassembling to a human operator and a cobot. Therefore, an action is the next component for its disassembly. The action space also consists of all available components for disassembling at each time step. Moreover, agents should take a new action (component) in each time step as each previously selected component (action) was disassembled from the product.

Reward

Several prominent features in a product disassembling represent the process's performance. Consequently, the main goal of a disassembly planning model is to optimize the process based on these features. In this research, we consider the operation time and the recovered part quality as the most impactful features of a disassembly process. Therefore, we define the reward function of the model based on these two features. Equation 4.7 illustrates the reward function, which is a linear combination of three variables. The first is the quality reward r_q, presenting the quality of a recovered component. The weight α_q also defines its corresponding impact on the reward function. The second variable is the time reward r_t that denotes the operation time. It is a penalty term to incentivize the agents for completing their tasks in a shorter time. We give this reward the weight α_t to see how it affects the reward function. Notably, the minus sign preceding α_t implies that a longer process time leads to less reward value and a punishment for the model, and vice versa.

In addition to the time and the quality rewards, we add the penalized reward r_p to the reward function. It aims to motivate the robot agent and the human agent to select easy tasks and difficult or very difficult tasks, respectively. In the case of the robot agent, if it chooses an easy task, the penalized reward will be positive. Also, it will be negative whenever the robot agent picks a difficult or very difficult task. Conversely, choosing easy tasks results in a negative value and punishment for the acceptable, medium, and high-skilled human agents incentivizing them to leave these tasks for the cobot. Thus, the value will be positive when they select difficult or very difficult tasks. These positive and negative values are discussed in more detail in Section 4.5.1. In addition, in the same way as two other coefficient terms, α_p indicates the impact of the penalized reward. Given that these variables are measured in different units, the coefficients also standardize their scales before summation.

$$r_{total} = \alpha_q \cdot r_q - \alpha_t \cdot r_t + \alpha_p \cdot r_p \tag{4.7}$$

4.4.4 Switching between the collaborative and inactive modes

During a disassembly process, sometimes only difficult and very difficult tasks are available. In these cases, performing a task by the cobot can damage its structure and a product's architecture. In order to prevent these issues, the cobot should stop working and wait for the human operator to complete the available tasks.

As an innovation, we propose an algorithm to switch the model from the multi-agent to the single-agent, in which a human operator works alone. In other words, we change the

human-robot interaction from the collaboration mode to the inactive mode. This switching process is shown in Figure 4.3. Moreover, Algorithm 2 presents the switching process in more detail. First, we select a node with the lowest difficulty score among all nodes provided by Algorithm 1. If its difficulty score is in the easy level, there is at least one node with the easy level. Therefore, the cobot can perform a task, and the model works in collaborative mode. Otherwise, the cobot stops working, and the model operates in inactive mode.

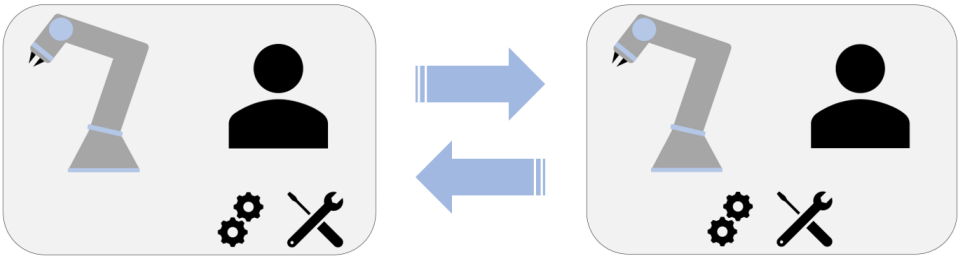


Figure 4.3 Switching between the collaborative and inactive modes

Algorithm 2 Switching between the collaborative and inactive modes

Check all available nodes with respect to the Algorithm 1
Select the node with the lowest difficulty score.
if the node's difficulty score is in easy level **then**
    Work in the collaborative mode.
else
    Work in the inactive mode.
end if

4.4.5 Real-time task allocation process

Several incidents may occur that deviates a disassembly process from its predetermined routine. Therefore, generating a task sequence before starting the process cannot address these unforeseen scenarios. Consequently, real-time task allocation is one of the most vital features of a disassembly planning model to cope with uncertainties.

This research proposes a model for allocating tasks in real time as an innovative alternative. The real-time feature requires that the model makes a decision at each time based on the product's conditions and operators' availability. Most previous works in disassembly planning theoretically determine an optimal task sequence before starting the process. Alternatively, our proposed model allocates tasks in real time, and does not consider planned times for tasks' completion.

Figure 4.4 illustrates the proposed real-time task allocation mechanism. At t = 0, the model assigns tasks 1 and 2 to the human operator and the cobot, respectively. Task 3 is allocated to the human operator when task 1 is performed at t = T_1. Similarly, the model recommends task 4 to the human operator after task 3 is completed at t = T_1 + T_3. Following this, the model assigns task 5 to the cobot when task 2 is completed at t = T_2. This process will continue until all tasks are completed. Section 4.4.6 explains this real-time task allocation mechanism in more detail.

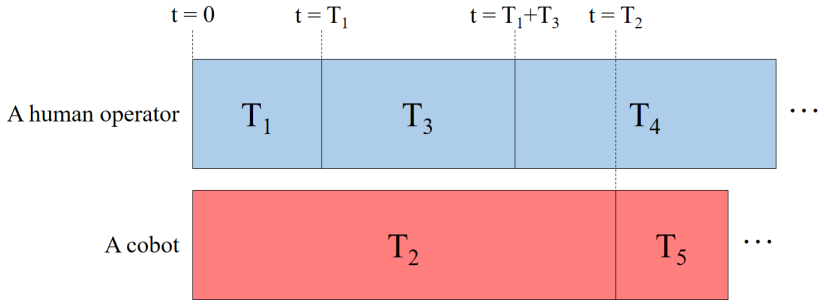


Figure 4.4 The real-time task allocation mechanism

4.4.6 A multi-agent deep reinforcement learning approach for disassembly planning

Deep Q-Network

We employ a multi-agent deep Q-network (DQN) architecture to model the proposed disassembly planning framework in Section 4.4.3. DQN is a popular reinforcement learning model introduced in [124]. It utilises two deep neural networks, prediction and target networks, to approximate Q-values. The prediction network is used to estimate Q(s,a), the Q-values at the current state for all possible actions. Additionally, the next state Q-values, denoted as Q(s', a'), are computed through the target network.

DQN aims to reach the optimal Q function, represented by Equation 4.8. It results in the maximum amount of cumulative rewards in the future steps by selecting an action a_t at state s_t regarding the policy π. A discount factor called γ is also a trade-off between the immediate and the long-term rewards. Equation 4.9 is also used to recursively update the Q-values at each step, in which the term α is the learning rate of the process.

$$Q^*(s,a) = \max_{\pi} \mathbb{E}\left[R_t + \gamma R_{t+1} + \gamma^2 R_{t+2} + \ldots \mid s_t = s, a_t = a, \pi\right] \tag{4.8}$$

$$\mathrm{Q(s,a)} \leftarrow \mathrm{Q(s,a)} + \alpha[\,\mathrm{R(s,a)} + \gamma \arg\max_{\mathrm{a'}} \mathrm{Q(s',a')} - \mathrm{Q(s,a)}\,] \tag{4.9}$$

Ultimately, the loss function of the process is shown in Equation 4.10, where θ_i and θ_i^- represent parameters of Q-values in the current and future states. Instead of computing the entire summation, stochastic gradient descent is normally used to update the prediction network parameters. Also, these parameters are loaded into the target network's parameters after each T_{target} time.

$$\nabla_{\theta_i} L\left(\theta_i\right) = \mathbb{E}_{(s,a,r,s')}\left[\left(r + \gamma \max_{a'} Q\left(s', a'; \theta_i^-\right) - Q\left(s, a; \theta_i\right)\right) \nabla_{\theta_i} Q\left(s, a; \theta_i\right)\right] \tag{4.10}$$

Real-time interactive human-robot collaboration disassembly planning model

This section proposes a DQN-based model to allocate tasks to agents in real time. The proposed framework and the associated algorithm with more details are displayed in Figure 4.5 and Algorithm 3, respectively. We initialize Θ_{h} and Θ_{r}, the parameters of the human agent and the robot agent, with random variables. We choose actions $\mathrm{a_t^h}$ and $\mathrm{a_t^r}$ at time t for the human operator and the cobot in the first step of starting an episode. This action selection process is performed with respect to Algorithms 1 and 2, as well as the ϵ-greedy strategy. This strategy balances the exploitation-exploration dilemma by the parameter ϵ.

In contrast to the previous models that assign tasks based on predetermined operation times, we allocate a new task to the human operator or the cobot once they complete the previous tasks. In this regard, we indicate a human operator and a cobot task completion times by $\mathrm{t'}$ and $\mathrm{t''}$. Whenever the human agent or the robot agent completes their corresponding tasks (at $\mathrm{t==t'}$ and $\mathrm{t==t''}$), they will receive rewards $\mathrm{r_{t'}^h}$ or $\mathrm{r_{t''}^r}$, respectively. In the case of the human agent, it moves from the current state $\mathrm{s_t^h}$ to the new state $\mathrm{s_{t'}^h}$. Then, the values $\mathrm{s_t^h}$, $\mathrm{a_t^h}$, $\mathrm{r_{t'}^h}$, and $\mathrm{s_{t'}^h}$ are stored in the replay memory $\mathrm{M_h}$. Next, we sample a batch from the replay memory $\mathrm{M_h}$ to train the human agent[1]. In this dynamic process, the human agent is updated at each step. Meanwhile, the robot agent may change the product's condition by disassembling a component. Therefore, we update the state $\mathrm{s_{t'}^h}$ based on these possible changes. This is a crucial step in the real-time task allocation process to ensure synchronous operation between agents. Consequently, completed tasks may be reassigned if this step is not performed. Finally, a new action $\mathrm{a_t'}$ is allocated to the human agent with respect to limitations presented in Algorithms 1 and 2, as well as the ϵ-greedy strategy. It is notable that the DQN output is the selected action. As a result, wherever an action is chosen, the

[1]For further details, please read [124]

DQN is employed.

We consider the mentioned steps for the cobot too, in which all values s_t^r, a_t^r, $r_{t''}^r$, and $s_{t''}^r$ are stored in the replay memory M_r. This algorithm delivers the learned human agent and robot agent models that can make decisions in real time. Furthermore, Table 4.3 illustrates all the key symbols of the proposed algorithm with their descriptions.

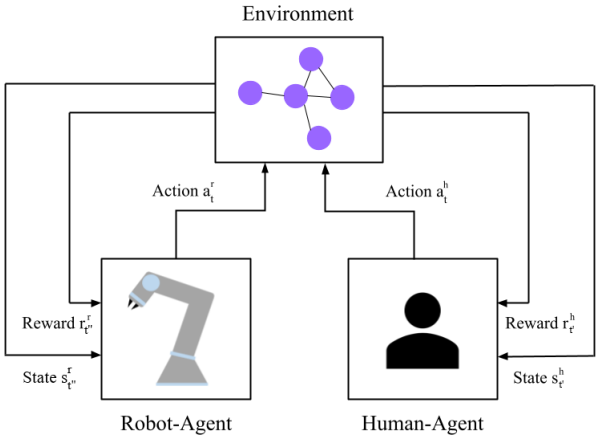


Figure 4.5 The proposed reinforcement learning-based framework

4.5 Experiments

In this section, the case study, experimental set-up, and results are discussed. First, the problem context, the evaluation protocol, and a simulation approach for environment modeling are presented. Then, we conduct three human-robot collaboration scenarios and perform sensitivity analyses based on the parameter ϵ, the discount factor γ, and the replay memory size, which represent the influential parameters of the model. Instead of serving as hyper-

Table 4.3 Description list of key symbols

Symbols	Description
γ	discount factor
ϵ	epsilon-greedy parameter
M_h	human agent replay memory
M_r	robot agent replay memory
Θ_h	human agent model's parameters
Θ_r	robot agent model's parameters
t'	A human operator's time required to complete the corresponding task
t''	A robot's time required to complete the corresponding task
s_t^h, a_t^h, r_t^h	state, action, and reward of the human agent at time step t
s_t^r, a_t^r, r_t^r	state, action, and reward of the robot agent at time step t
T_{target}	The number of required iterations to update the target network
$\pi^\star$	Optimal policy
α	Learning rate

Algorithm 3 Human-robot collaboration disassembly planning framework

Input: the parameter ϵ, T_{target}, the product's graph, replay memories M_h and M_r the number of time steps T, the number of episodes N, mini-batch size, discount factor γ

Initialize the human agent and robot agent models' parameters Θ_h and Θ_r randomly.
for Episode n = 1,...,N **do**
Select an action a_t^h in s_t^h by the human agent based on the limitations in Algorithm 1, Algorithm 2, and ϵ.
Select an action a_t^r in s_t^r by the robot agent based on the limitations in Algorithm 1, Algorithm 2, and ϵ.
for Time step t = 1,...,T **do**
if $t = t'$ **then**
Receive the reward $r_{t'}^h$
State s_t^h evolves to new state $s_{t'}^h$
Store the tuple $(s_t^h, a_t^h, r_{t'}^h, s_{t'}^h)$ in the replay memory M_h.
Update the human agent's parameters Θ_h by sampling a batch from the replay memory M_h.
Update $s_{t'}^h$ based on the possible changes in the product condition caused by the robot agent.
Select an action $a_{t'}^h$ in $s_{t'}^h$ by the human agent based on the limitations in Algorithm 1, Algorithm 2, and ϵ.
end if
if $t = t''$ **then**
Receive the reward $r_{t''}^r$
State s_t^r evolves to new state $s_{t''}^r$
Store the tuple $(s_t^r, a_t^r, r_{t''}^r, s_{t''}^r)$ in the replay memory M_r.
Update the robot agent's parameters Θ_r by sampling a batch from the replay memory M_r.
Update $s_{t''}^r$ based on the possible changes in the product condition caused by the human agent.
Select an action $a_{t''}^r$ in $s_{t''}^r$ by the robot agent based on the limitations in Algorithm 1, Algorithm 2, and ϵ.
end if
end for
end for

Output: The trained human agent and robot agent models,

parameter tuning, these experiments aim to assess the model's robustness and adaptability under different operational scenarios.

4.5.1 Experimental setup

Data and evaluation protocol

In this study, we exploit a hard disk drive presented in [1] as a case study to evaluate our proposed model. The hard disk drive is a type of waste electrical and electronic equipment (WEEE). WEEE is on the rise globally with a growing rate. This amount of waste is generated by appliances at home, information technology, and telecommunications equipment. As one of the fastest-growing waste streams, rapid technological advancements, consumer demand for the latest electronics products, and difficulties in repairing them result in this amount of waste [125]. The use of remanufacturing techniques, such as disassembly, enables companies to sustainably reuse or recycle WEEE, resulting in several positive impacts on business and the environment. The increasing amount of WEEE results in rising demands for disassembly. Hence, the classical manual disassembly method cannot meet the extreme demands of industries. Consequently, there is a need to switch from manual disassembly to robotization disassembly. Notably, robots cannot perform all tasks due to their lack of flexibility. Hence, implementing disassembly collaborative cells, where human operators work alongside cobots, is necessary. As a result, both the flexibility of humans and the efficiency of cobots are involved in disassembly.

Figure 4.6 depicts the graph of the product that consists of 15 components. As discussed in Section 4.4.3, we integrate several physical and geometric features of a product's components into our model. Most of these features, such as size, weight, accessibility, and liaison types, should be outlined by the product manufacturers. Table 4.4 illustrates all the mentioned features and their related values.

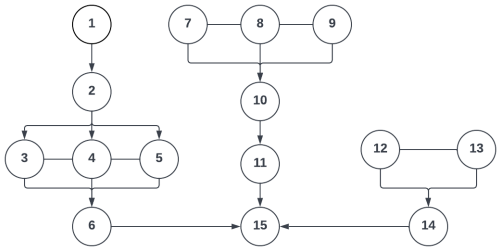


Figure 4.6 The product structure presented in [1]

To assess the performance of our model, we use the cumulative reward of each episode, which

is a summation of the reward values in T number of time steps. Equation 4.11 illustrates the mentioned formula, where r_t and R are the reward at time step t and the cumulative reward, respectively.

$$R = \sum_{t=1}^{T} r_t \tag{4.11}$$

Environment

We designed a mathematical rule-based algorithm as the environment of the reinforcement learning framework and its corresponding feedback. Although we train our proposed model using feedback from this rule-based environment, the delivered model, due to its dynamic nature, can make decisions in unexpected states of the environment in real-world scenarios. While many traditional sequential optimization models fail in changing environments due to uncertainties.

We categorize the difficulty scores into the easy, difficult, and very difficult levels by $Threshold_1$ and $Threshold_2$, which are defined experimentally. We also assume that human operators of varying skill levels perform disassembly tasks of varying difficulty, with varying disassembly times and recovered part quality. In this case, even though human operators with an acceptable level of experience and skill can perform easy tasks with an appropriate performance at an appropriate time, they will lose their performance in the face of difficult and very difficult tasks. Human operators with medium-skill, on the other hand, can only do simple tasks efficiently. They also finish difficult tasks in a reasonable time and with acceptable quality. In the event of very difficult tasks, they can only complete them by attaining an acceptable quality in an inefficient time. High-skilled human operators, on the other hand, can complete

Table 4.4 Disassembly features and their corresponding values

Component	Strength	Weight	Size	Destructive	Positioning	Accessibility	Liaisons scores	Liaisons properties
1	4	2.4	3.5	1	5	1	10	2
2	1	2	2	0	1.2	1.6	2	1.2
3	2	2	2	0	1.2	1	3	1.3
4	1	2.2	4	0	2	2	2	1.4
5	2	2.2	3.5	0	1.2	1.6	4	1
6	1	2	2	0	5	1	2	1.1
7	4	2.4	3.5	1	2	1	6	1.4
8	2	2	2	0	1.2	2	3	1.2
9	2	2.2	2	0	2	1.6	2	1.1
10	1	2	2	0	2	1	4	1.5
11	1	2.2	2	0	1.2	1.6	2	1.7
12	4	4	4	1	1.2	1	10	1.8
13	1	2	2	0	1.2	2	2	1.1
14	2	2.2	2	0	1.2	2	3	1.9
15	2	4	4	0	5	2	4	1.5

tasks of varying complexity levels effectively.

Cobots can only perform simple and repetitive operations. Therefore, easy tasks should be assigned to a cobot. Cobots cannot complete difficult and very difficult tasks due to their lack of physical capacity, and it is necessary to learn the model to prevent allocating these tasks to cobots.

The mathematical modeling of the environment based on rules is shown in more detail in algorithm 4. The first condition describes a scenario in which an experienced, acceptable-level human operator collaborates on disassembly. Accordingly, he/she performs the task appropriately only when it is easy. The second and third conditions are demonstrated situations where medium- and high-skilled human operators are working with a cobot, respectively. While a medium-skilled human operator can only complete easy tasks with high expertise, a highly experienced human operator can perform tasks of all difficulty levels efficiently. Finally, condition four shows a robot that can only perform easy tasks competently.

As mentioned in Section 4.4.3, we import the recovered part quality and disassembly time as two feedbacks from the environment. To train the model, we consider discrete values for each level of the recovered part quality and time. In terms of the recovered part quality, the values 2, 7, and 10 are selected for the inefficient, appropriate, and efficient quality, respectively. Additionally, we consider the values of 1, 2, and 3 for the efficient, appropriate, and inefficient time levels. It should be noted that we use these values to reduce the time of our simulation as proof of concept.

Furthermore, to consider the penalized reward introduced in Section 4.4.3, the values 10 and -10 are utilised for the positive and negative rewards, respectively. Moreover, we consider the value 15 as a penalized reward for highly skilled human operators performing extremely difficult tasks. The goal is to motivate the operator to choose very difficult tasks that other types of operators cannot complete efficiently. All the mentioned values for human operators' type and a cobot at each task difficulty level are shown in Table 4.5. As a cobot cannot perform tasks at difficult and very difficult levels, there is no value in their corresponding cells in Table 4.5.

4.5.2 Results and discussion

We proposed a context-aware reinforcement learning-based architecture to model human-robot disassembly task allocation processes. As it aims to consider human operators' skills in the decision-making process, we evaluate and validate our model in three scenarios containing a cobot's collaboration with human operators in different skill levels: acceptable, medium,

Algorithm 4 The environment's mathematical rule-based modeling

if The human operator has an acceptable-level of experience **then**
  if The task's difficulty < Threshold_1 **then**
    Appropriate recovered quality at an appropriate time.
  else
    Accepted recovered quality at inefficient time.
  end if

else if The human operator has a medium level of experience **then**
  if The task's difficulty < Threshold_1 **then**
    Efficient recovered quality at efficient time.
  else if Threshold_1 <= The task's difficulty < Threshold_2 **then**
    Appropriate recovered quality at an appropriate time.
  else
    Acceptable recovered quality at inefficient time.
  end if

else if The human operator has a high level of experience **then**
  Efficient recovered quality at efficient time.

else The operator is a cobot
  if The task's difficulty < Threshold_1 **then**
    Efficient recovered quality at Efficient time
  else
    Not completed.
  end if
end if

Table 4.5 Time and recovered part quality of disassembly tasks in each difficulty level based on operators' skill level

Index	**Easy**			**Difficult**			**Very difficult**		
Operators	r_t	r_q	r_p	r_t	r_q	r_p	r_t	r_q	r_p
Acceptable-skilled human	2	7	-10	3	2	10	3	2	10
Medium-skilled human	1	10	-10	2	7	10	3	2	10
High-skilled human	1	10	-10	1	10	10	1	10	15
Cobot	1	10	10	Not completed		-20	Not completed		-20

and high. For this purpose, we use three human agents, and each of them plays the role of a human operator with a particular skill level.

Furthermore, the proposed architecture has several important parameters that can have a considerable impact on the final decision. In this part, we discuss and analyze the results of the proposed framework with different parameter values to determine the best values for them in each scenario. In this regard, we examine the model's sensitivity to the parameter ϵ, the discount factor γ, and replay memory size. To begin with, we use fixed values for the parameter γ and replay memory size to evaluate the model's performance while changing the values of the parameter ϵ in a diverse range to select the optimal value. After finding the optimum value, we repeat the same procedure for the parameters γ and replay memory size, in that order. The order of parameter configuration is determined based on the importance of each parameter in the model's learning process.

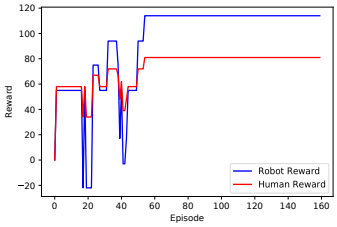
In addition, we consider the corresponding weights and coefficients of the difficulty vector and the reward terms in Equations 4.5 and 4.7 as values 1, presenting a standard situation.

Scenario I: Collaboration of a cobot and an acceptable-skill level human operator

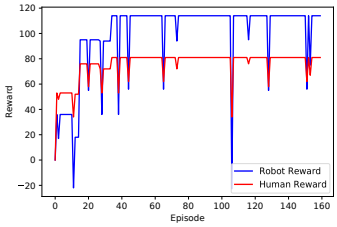
Sub-figures 4.7a, 4.7b, and 4.7c illustrate the model's reward plots per episode for three values of the parameter ϵ: 0.0, 0.01, and 0.02. The observations indicate that the higher values, leading to more exploration by the agents, result in greater fluctuations in the reward plots. In this scenario, the agents obtain better outcomes by exclusively operating in the exploitation mode. They therefore achieve the best performance by the value 0.0, where they do not explore.

As shown in Sub-figures 4.7d to 4.7i, we analyze the model's performance with six different values of the parameter γ. The best result among all values is achieved by the value 0.2, which is shown in Sub-figure 4.7i.

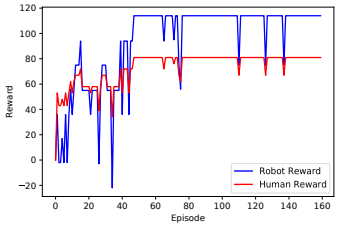
Replay memory is a buffer that stores the last number of N transitions during the interactive process between an agent and an environment. These transitions are sampled in each iteration



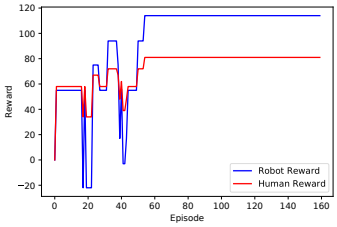
(a) $\epsilon = 0.0$



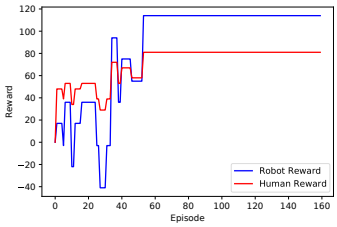
(b) $\epsilon = 0.01$



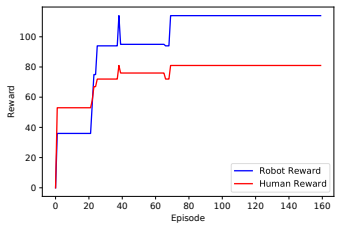
(c) $\epsilon = 0.02$



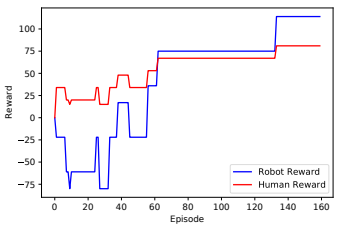
(d) $\gamma = 0.0001$



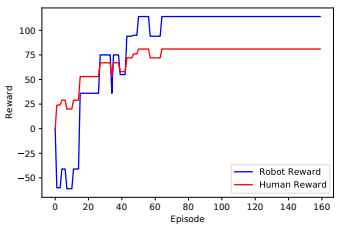
(e) $\gamma = 0.01$



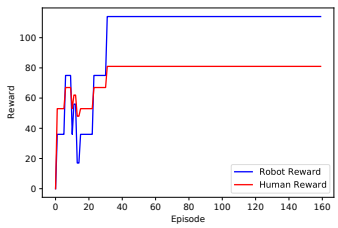
(f) $\gamma = 0.05$



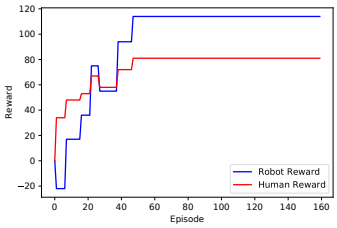
(g) $\gamma = 0.1$



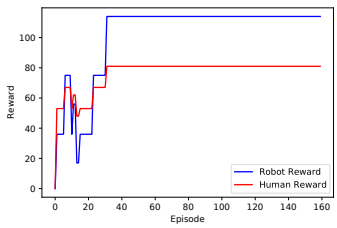
(h) $\gamma = 0.15$



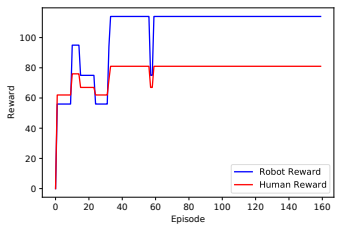
(i) $\gamma = 0.2$



(j) Replay memory size = 50



(k) Replay memory size = 100



(l) Replay memory size = 200

Figure 4.7 Sensitivity analyses in Scenario I

to train DQN. Moreover, replay memory size can be a significant factor in the learning process. To select an optimal replay memory size, we evaluate the model by three sizes: 50, 100, and 200. The reward plots per episode for these three sizes are illustrated in Sub-figures 4.7j, 4.7k, and 4.7l. Here, the model obtains a more robust result by replay memory size 100. Notably, the sizes 50 and 200 result in more fluctuations in the reward plots. In the case of the size 200, we infer that storing a higher number of transitions in the replay memories leads to a decline in the sampling performance. On the other hand, we conclude that size 50 is not enough for the replay memories, and it reduces the performance of the learning process.

Scenario II: Collaboration of a cobot and a medium-skill level human operator

Sub-figures 4.8a, 4.8b, and 4.8c display the model's performance by different values of the parameter ϵ. Similar to Scenario I, the model achieves a more robust result by the value 0.0 than other values.

The reward plots per episode for six parameter γ values are shown in Sub-figures 4.9d to 4.9i, where the model by the value 0.0001 converges faster than other cases. In this scenario, we can infer that the agents work more efficiently by focusing on the current Q-values instead of considering the future Q-values.

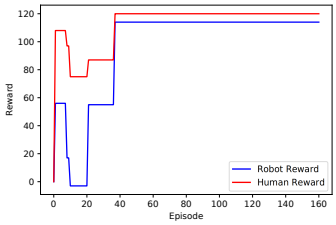
In terms of replay memory size, Sub-figures 4.8j, 4.8k, and 4.8l show the evaluation of the model under three sizes. We can conclude the same inference of Scenario I for replay memory size as the model outperforms by the size 100 than other sizes.

Scenario III: Collaboration of a cobot and a high-skill level human operator

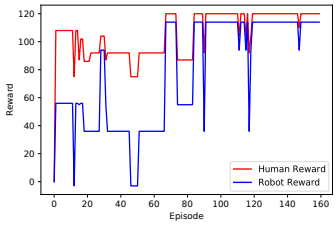
Similar to Scenarios I and II, the agents achieve better results without performing in the exploration phase. As shown in Sub-figures 4.9a, 4.9b, and 4.9c, the parameter ϵ value 0.0 provides the most stable and fast-convergence result.

Sub-figures 4.9d to 4.9i show the model's evaluation by different values of the discount factor γ. Here, the model converges faster by the value 0.15 compared to the other values.

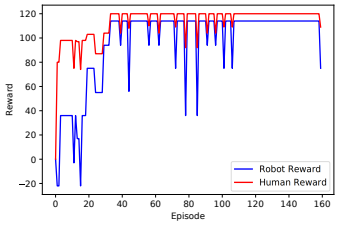
Experimental analysis in Sub-figures 4.9j, 4.9k, and 4.9l demonstrates that the replay memory size 100 has the most stable and fastest convergence result. Therefore, we can infer the same conclusion of Scenario I in the terms of replay memory size in this scenario.



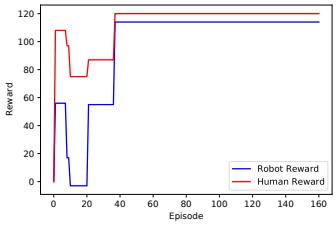
(a) $\epsilon = 0.0$



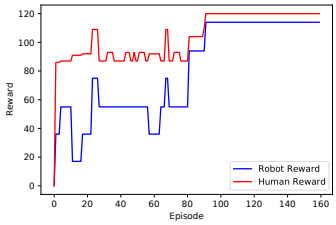
(b) $\epsilon = 0.01$



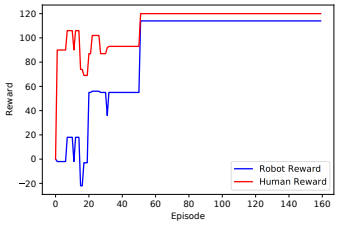
(c) $\epsilon = 0.02$



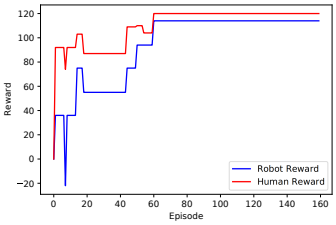
(d) $\gamma = 0.0001$



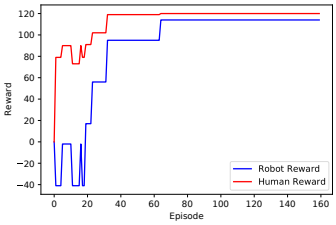
(e) $\gamma = 0.01$



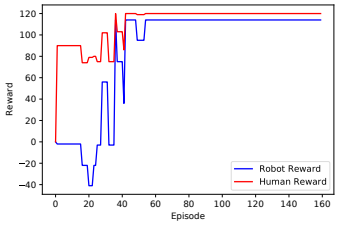
(f) $\gamma = 0.05$



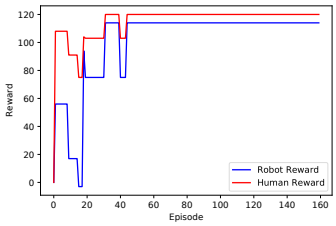
(g) $\gamma = 0.1$



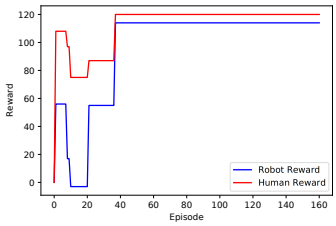
(h) $\gamma = 0.15$



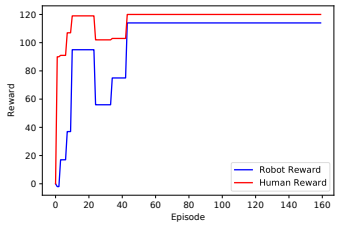
(i) $\gamma = 0.2$



(j) Replay memory size = 50

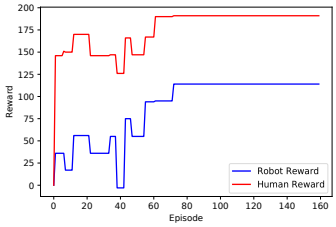


(k) Replay memory size = 100

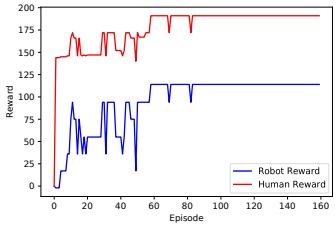


(l) Replay memory size = 200

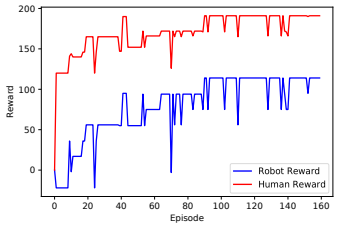
Figure 4.8 Sensitivity analyses in Scenario II



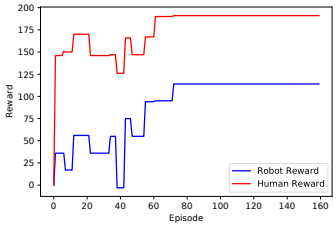
(a) $\epsilon = 0.0$



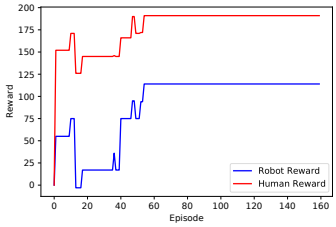
(b) $\epsilon = 0.01$



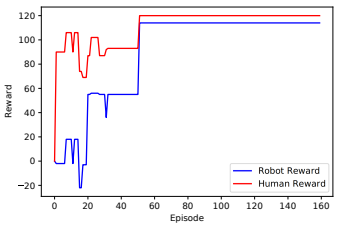
(c) $\epsilon = 0.02$



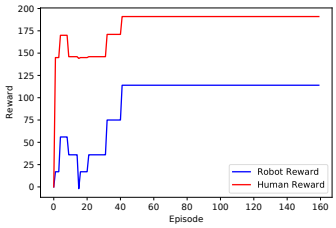
(d) $\gamma = 0.0001$



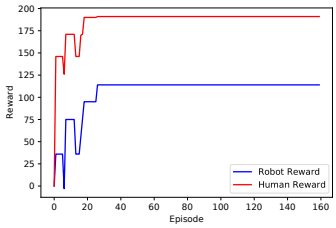
(e) $\gamma = 0.01$



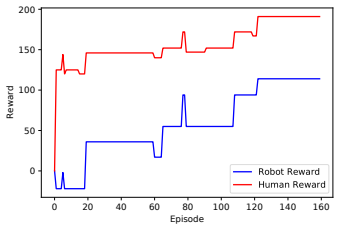
(f) $\gamma = 0.05$



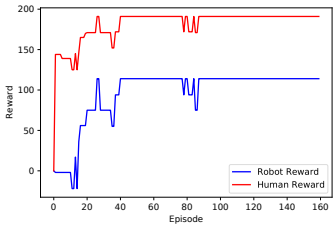
(g) $\gamma = 0.1$



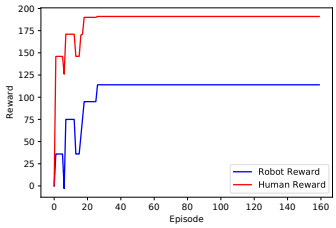
(h) $\gamma = 0.15$



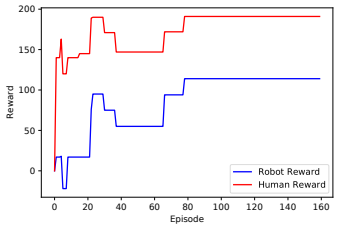
(i) $\gamma = 0.2$



(j) Replay memory size = 50



(k) Replay memory size = 100



(l) Replay memory size = 200

Figure 4.9 Sensitivity analyses in Scenario III

Comparative evaluation

The comparative evaluation provided in this part is conducted in terms of the time required for convergence. The comparative evaluation provided in this part is conducted in both quantitative and qualitative metrics. We utilise the time required for convergence as a quantitative metric. In this regard, we compare the convergence time between our model and the model presented in [1], where we utilised their case study. Although the required disassembly features and the evaluation protocol in these two models are different, we examined the number of iterations as a metric to assess the convergence time. The proposed model in [1] requires 120, 127, and 189 iterations for convergence based on their experimental analysis. In contrast, the convergence times for different scenarios in our model are 37, 26, and 31, and we achieve significantly faster convergence. We know that the definition of iteration in two models are different, and we just used convergence time as key metric to highlight the relative efficiency of the model in finding a suitable disassembly plan. Hence, our proposed model performs better than the baselines in terms of the convergence time and required disassembly features, which were analyzed in Section 4.3.4.

In terms of quantitative metrics, we compare the disassembly features incorporated in our modeling to those used in recent literature, which we consider as the baselines. Figure 4.10 illustrates the baselines and their corresponding features. We consider real-time decision-making, difficulty, operation time, skill levels, recovered quality, self-adaptiveness, and human-robot collaboration during disassembly as the most significant features.

Although each baseline includes several features in modeling, none incorporates all the features. In contrast, we present a comprehensive disassembly planning model that integrates all the required features. Our model is designed to operate in a human-robot collaboration cell, where a cobot assists a human operator. Moreover, our model makes a decision in real time based on the tasks' difficulty, human skill levels, and availability of the human operator and the cobot. Furthermore, the main objectives of the model are to minimize operation time and maximize the quality of recovered parts. Notably, the model adapts itself in response to dynamic variations in the environment.

4.6 Conclusion

A disassembly process depends on how tasks are split between a human operator and a cobot in a cell. Assigning incorrect tasks to operators can result in longer processes and lower-quality outputs. Real-time decision-making is another essential aspect of a task allocation model. Many unexpected events can lead to changes in the process routine. Consequently,

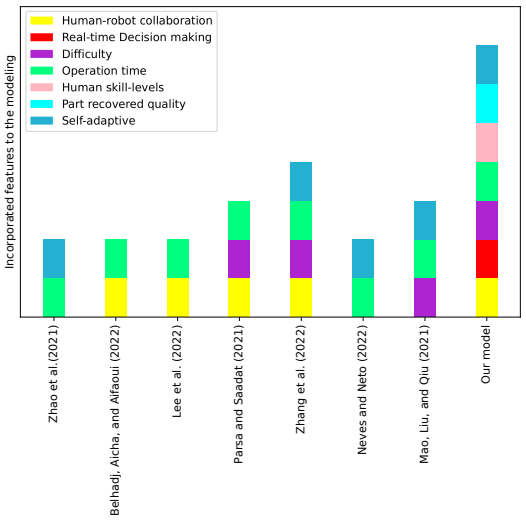


Figure 4.10 A qualitative comparison of recent literature

most optimization models that theoretically generate task sequences prior to the process are ineffective for making decisions in uncertain situations because they cannot change these predetermined task sequences in real time, which is crucial for adapting to unexpected events and optimizing performance in uncertain situations.

This paper focuses on human-robot collaboration in disassembly processes by proposing a multi-agent reinforcement learning-based planning model. The model consists of human and robot agents, where each agent selects tasks to be assigned to its corresponding operator. Instead of utilizing a preplanned sequence, the agents select tasks in real time based on the current situation of the human operator, cobot, and product. In our analysis, we considered three types of human operators with different work efficiencies according to their skill levels.

In terms of the product characteristics, we used several significant features of the product's components to compute the difficulty scores for the disassembly of each component. We introduced a graph-based approach as the EoL product architecture to categorize the product's components into different levels according to the bill of materials. The model involves these levels in the task allocation process. However, the aforementioned process may sometimes be challenging. In such cases, assigning a difficult task to a cobot may negatively affect its physical structure. To address this, we added the ability of the model to switch from a multi-agent mode to a single-agent mode. Tasks were also assigned to the human agent, and the robot agent waited to find the appropriate task.

We evaluated our model under three scenarios in which human operators of different skill levels collaborated with a cobot. The experimental results demonstrated that the proposed

model outperforms the model presented in [1]. Although we validated the model in scenarios in which a human operator collaborated with a cobot, its application can be extended to scenarios where multiple human operators collaborate with multiple cobots. Further investigations can focus on examining other features of human-robot collaboration in a disassembly process, such as cobot energy consumption and human safety. In addition, further research could investigate the relationship between difficulty and the effective features. In the next step of this research study, this relation will be evaluated and assessed in an experimental analysis. Our goal is to design a fuzzy-based model to analyze features' interactions for estimating a more precise relationship, similar to the approach presented in [126].

Disclosure statement

No potential conflict of interest was reported by the authors.

Data availability statement

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Funding

The authors gratefully acknowledge the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) for this research project [grant number RGPIN-2020-05565].

CHAPTER 5 ARTICLE 2: REAL-TIME SUSTAINABLE COBOTIC DISASSEMBLY PLANNING USING FUZZY REINFORCEMENT LEARNING

Authors: Ashkan Amirnia and Samira Keivanpour

Submitted on March 8, 2024, and published online on November 26, 2024, in *International Journal of Production Research (IJPR)*, pages 1-24, DOI: 10.1080/00207543.2024.2431172.

5.1 Abstract

Collaborative robots (cobots) play a vital role in smart manufacturing, particularly in disassembly processes. Human-robot collaboration (HRC) methods simultaneously leverage the complementary capabilities of humans and cobots, offering promising improvements in disassembly processes. A review of the literature reveals that most proposed HRC disassembly planning models do not incorporate sustainable factors, such as consumed energy, human safety, ergonomic risks, and circularity, in the decision-making process. Furthermore, uncertainties inherent in disassembly processes, such as the quality of recovered parts, are not well-addressed in the literature. This paper presents a novel multi-agent fuzzy reinforcement learning (RL) planning model for sustainable HRC disassembly. In addition to cost elements, the developed model involves social and environmental considerations in the real time planning process. By developing a fuzzy-based environment in the RL architecture, the proposed approach aims to effectively model the possibilistic uncertain parameters involved in the problem.

Experimental analysis shows that the model presented in this research outperforms a baseline model, applied to the same case study, in terms of convergence time. Furthermore, in terms of qualitative analysis, the proposed model integrates a more extensive set of features into the planning process compared to recent literature.

Keywords: Human-robot collaboration; disassembly planning; reinforcement learning; sustainable driven planning; fuzzy logic; cobotic disassembly

5.2 Introduction

Nowadays, many developed and developing countries face the increasing production of end-of-life (EoL) waste as one of the most common environmental issues. Various methods

exist for handling this type of waste. The waste management hierarchy pyramid explains several methods including prevent, reduce, reuse, recycle, landfill, and disposal for dealing with waste. The primary objective of the prevent and reduce approaches is to mitigate waste generation by addressing several factors, such as product and process design, and the operation phase. Moreover, the reuse methods focus on utilizing a product multiple times or repurposing its parts in remanufacturing processes. This strategy is more sustainable than recycling, landfill, and disposal methods, which cause substantial environmental costs.

Product disassembly addresses the process of extracting parts of an EoL product, facilitating their reuse in remanufacturing processes for new fabrication. Consequently, fewer raw materials are consumed in production processes, resulting in considerable economic and environmental gains. Thus, product disassembly contributes a critical function in circular economy and remanufacturing. Disassembly planning is a key step in these processes, and it refers to efficiently generating a task sequence aimed at minimizing costs and maximizing profits. An efficient disassembly approach helps the manufacturing industry to decrease negative environmental and economic impacts [64]. This research uses the terms "product" and "task" to refer to an EoL product that is targeted for disassembly and the process of disassembling a component from a product, respectively.

The involvement of cobots as human collaborators in industrial processes, such as product disassembly, is increasing. These cobots can efficiently complete repetitive and simple tasks that human operators may not perform accurately due to fatigue or distraction. Additionally, they can undertake tasks that pose dangers to human health or may cause long-term muscle disorders. However, many cobots lack the requisite power and flexibility to effectively complete all types of tasks. Therefore, the synergistic HRC approach efficiently gains from the flexibility and power of humans as well as the capability of cobots to precisely carry out repetitive and dangerous tasks. Several researchers have proposed different data-driven approaches to plan HRC disassembly processes. [127] introduces a hybrid HRC disassembly planning model based on a particle swarm optimization (PSO) and the Q-learning algorithm. In each evolutionary step, the Q-learning approach aims to choose the most appropriate optimizer to enhance the PSO algorithm performance. The proposed model targets to allocate tasks including toxic material to cobots, as well as assign tasks dealing with challenging parts to humans. The authors also applied several perturbation approaches and local search techniques (tabu search, cooperative local search, variable neighborhood search, insert, inverse, and swap) to the framework for preventing the model from falling at local minimums. [88] presents an HRC disassembly planning model, in which the authors initially extracted all precedence constraints from CAD data of a gearbox reducer. In the next step, they generated all feasible task sequences according to the precedence relations. Following that, the

sequence with minimum disassembly operation time is chosen. Herein, the authors address the disassembly time of each part as a predetermined constant value. Guo et al. [128] have proposed a planning method for destructive HRC disassembly processes. By considering failure characteristics of products, the method computes disassembly modes along with the associated qualities of recovered parts. Additionally, it obtains the physical limitations and precedence relations of parts to form a product constraint model. The process is optimized using an extended genetic algorithm (GA), considering factors such as task difficulty, operation time, cost, physical constraints, and product failure criteria. In another research study, [86] introduces an HRC disassembly planning model that represents the structure of an EoL product and the associated precedence relations by using an AND/OR graph. The proposed approach classifies tasks into eight different groups according to their corresponding difficulty and complexity. It optimizes the process by applying GA, with the aim of allocating more challenging and heavy tasks to the human operator, while assigning easy and repetitive tasks to the cobot. The corresponding fitness function of GA comprises several objectives, such as indices of untargeted components, operation time, and the frequency of human operator change. [89] introduces an HRC disassembly planning model based on a cascade structure. With the help of a digital twin framework, the model fuses data from real and virtual spaces. Then, it sequentially preprocesses the data to obtain parts features, types of liaison and tools, as well as disassembly precedence constraints. The model determines the most optimal sequence by using GA.

According to [2] and [48], because of several factors, such as extreme intensity and frequency of use, the structure of a product may deviate from its standard configuration. Furthermore, the prolongation or failure of tasks deviates a decision-making process from its ideal routine. Therefore, it is necessary to update the task sequence according to the current status of the disassembly process and applying predetermined task sequences is impractical. The majority of previous research developed fixed planning models that cannot adapt according to possible uncertainties of the process.

A few researchers have addressed sustainable aspects in HRC disassembly planning. Most previous research studies plan disassembly sequences without respect to the environmental and social aspects, such as consumed energy, parts circularity, and human safety. In addition, most of the existing planning methods cannot effectively represent uncertain parameters, including recovered parts quality and difficulty. This uncertainty in data has a possibilistic nature[1] and is not addressed appropriately in the literature.

This research study aims to fill the mentioned gaps by presenting the following contributions.

[1]For further details, please read [129]

- Propose a real-time sustainable HRC disassembly planning model based on an RL architecture. This model allocates tasks in real time according to cost factors as well as elements concerning environmental and social indicators, outlined in Section 1.2.
- Develop new states and reward functions for the RL model according to the sustainability elements in addition to the technical feasibility of tasks.
- Design a fuzzy-based environment for the proposed RL architecture to effectively cope with the mentioned uncertainty in data.

The remainder of this paper is structured as follows. Section 5.3 reviews the related works. Following that, Section 5.5 explains the methodology. Section 5.6 discusses the experimental settings and results. Lastly, Section 5.7 concludes the paper, coupled with some ideas for the future steps.

5.3 Related works

This section reviews the related literature. First, it explores the problem of product disassembly planning. It then discusses the collaboration of humans and cobots in assembly/disassembly processes. Next, it examines the sustainability elements integrated into this problem. Finally, it synthesizes the literature to analyze the strengths and weaknesses of previous works and highlight potential gaps.

5.3.1 Disassembly planning

There is a considerable amount of literature on planning disassembly processes. Researchers have widely used various rule-based and learning-based algorithms to generate optimum task sequences. Classical optimization algorithms such as GA, PSO, and ant colony optimization (ACO) have been commonly used in the literature for disassembly planning [2].

Many researchers addressed product disassembly planning using fuzzy logic methods. [130] develops a multi-objective fuzzy-based model for balancing disassembly lines by considering safety issues. A hybrid fuzzy-GA approach for parallel disassembly planning is developed in [131]. They validated the model's efficiency by applying it to a case study involving a hydraulic press. [132] proposes a new graph-based fuzzy approach for planning product disassembly. The author evaluated the proposed approach by applying it to a telephone as a real case study.

Due to the limitations in many classical optimization methods in coping with the uncertain nature of disassembly planning, learning-based methods are increasingly becoming popular.

As the disassembly planning problem cannot be formulated within the domains of supervised or unsupervised learning, RL methods are frequently employed to determine the optimal task sequence. Over the past few years, several researchers have developed disassembly planning approaches based on RL models. [46] introduces a matrix-based model based on a Q-learning algorithm to optimize a disassembly process according to parts physical constraints and operation time, which is a summation of basic disassembly time, tool changing time, required time for tool positioning, and cleanup time. [47] represents the disassembly process by an extended Petri net (EPN), including several parts features, such as revenue, and cost. A maximum likelihood approach is then employed to estimate uncertain parameters such as revenue and cost. The authors used a tabular Q-learning algorithm to plan the process. Here, the state is defined based on the EPN's token, while the reward function is the difference between revenue and cost. Allagui et al. [133] developed a disassembly planning approach based on RL. It begins by extracting CAD data from a product's 3D model and then builds a matrix to generate all possible disassembly sequences in the x, y, and z axes. Furthermore, the proposed approach forms a reward matrix based on the physical feasibility of moving toward axes and a fitness function, considering the volumes of parts, operation time, wear parts, and changes in the process direction and tool. Lastly, a Q-table is defined to obtain all state-action Q-values. [48] introduces an RL disassembly planning model based on a graph approach, which represents the product structure by dividing it into layers based on geometry. Task allocation begins from the outer layers and progresses inward. The reward function compromises operation time, profit, and penalties for incomplete tasks before moving to the next inner layer. Next, [45] presents a hybrid GA-RL model to optimize disassembly sequences. They also used a Petri network to represent the process.

5.3.2 Human-robot collaboration

HRC in assembly/disassembly planning has received much attention in recent years, and many researchers have examined this subject. [134] introduces an HRC disassembly line balancing model. The model addresses the distance between humans and cobots as a safety factor, accounting for variable cobot speeds based on proximity to humans. The model employs an AND/OR graph representing a product's structure and constraint relationships, with a decision tree algorithm classifying tasks by several criteria, such as component weight and required tool type. An enhanced discrete Bees algorithm optimizes the process, aiming to reduce the number of workstations, operation time, and demand index. In another research study, Zhang et al. [92] developed an HRC assembly planning model based on a dual-agent RL architecture, in which an agent is responsible for task allocation to a human operator, while another agent selects tasks for a cobot. Furthermore, the authors presented an extended

version of the deep deterministic policy gradient (DDPG) model. It defines a global Q-value based on the combination of the two agents' Q-values, facilitating the synchronizing performance between agents.

Liu et al. [90] have proposed an HRC disassembly planning model based on a theoretical recursive learning algorithm. Initially, data from different sources is fused, extracting high-level information like human and cobot locations, hand gestures, body skeletons, and required tools. Subsequently, an RL model optimizes the process based on this information. Following that the operators execute the planned tasks. Finally, the acquired knowledge is shared using incremental learning and transfer learning techniques via cloud-based technology. In the scope of lithium-ion battery recycling, [91] proposes a heuristic-based approach to assess the resilience of HRC disassembly based on several factors, such as stability, redundancy, efficiency, and adaptation.

A comprehensive examination of recent works reveals that the use of cobots as human partners in assembly/disassembly processes is rapidly growing. It can be predicted that in the upcoming years, cobots will play a major role in the manufacturing industry.

5.3.3 Integrating sustainability criteria into disassembly planning

Manufacturing industries face great pressure to formulate sustainable development strategies due to existing environmental concerns, such as climate change, pollution, and depletion of resources [135]. The rising demand for sustainable development motivates companies to switch from conventional to sustainable manufacturing. It addresses the production of goods to deliver economic benefits while reducing environmental impacts and ensuring stakeholders' social responsibility during a product's life cycle [136]. Economic, environmental, and social are three major pillars of sustainable development.

A sustainable disassembly planning model optimizes the process based on economic, environmental, and social objectives. Several researchers have addressed sustainability in the context of disassembly planning. [137] proposes a disassembly line balancing model that sustainably allocates tasks based on economic, safety, and environmental criteria. Zhou and Bian [138] have developed a sustainable robotic disassembly line balancing model. To increase the robustness of the model in real-world scenarios, the authors considered tool change, uncertain operation time, and different robots in terms of efficiency and energy consumption. [139] develops an integrated sensor-based framework for sustainable product maintenance and disassembly.

The most important sustainability elements in this context are discussed below.

Energy consumption Sustained economic growth relies on energy, yet several factors, such as the increasing energy demand caused by global population growth, threaten modern economic progress [58]. The amount of consumed energy also plays a critical role in cobotic assembly/disassembly operations. Consuming more energy produces more energy footprints, which negatively impacts the environment. Moreover, uncontrolled energy consumption increases economic costs [2].

A few recent studies have paved the way to integrate consuming energy into the robotic disassembly planning process. [59] develops a robotic parallel disassembly sequence planning framework using the artificial bee colony (ABC) algorithm, aiming to minimize the makespan and total energy consumption of robots. They assumed the consumed energy arises in three stages: during disassembling, while waiting for new tasks, and during tool change. Similarly, [60] optimizes a robotic disassembly process to achieve maximum profit while minimizing energy footprint by employing the Bees algorithm. In the context of assembly/disassembly line balancing, [61] proposes a model based on an improved bi-objective evolutionary algorithm to optimize a robotization assembly line. As two main objectives, the model attempts to minimize total energy consumption and the number of workstations. In another study, [62] introduces a modified GA algorithm to solve the assembly line balancing and part feeding problem. The algorithm considers reducing energy consumption, the number of stations, and the number of supermarkets as the main objectives. [63] proposes a mixed model to minimize cycle time, total consumed energy, and peak workstation energy consumption in a robotic disassembly line balancing problem.

Despite the mentioned considerable efforts made in the planning of robotic disassembly, the aspect of energy in cobotic disassembly planning still remains largely unexplored by researchers. Hence, optimizing energy consumption of cobots in the disassembly process is an emerging research area.

Safety Human safety is a crucial concern for manufacturing companies. It is essential to avoid assigning dangerous operations to humans in assembly/disassembly processes. Many EoL products, such as waste electrical and electronic equipment (WEEE), include hazardous substances, which adversely affect human health and safety. It is imperative to keep away these materials from humans [64]. Human operators' safety is also a critical issue in HRC, and a growing body of literature has investigated this context. Mukherjee et al. [65] have presented a comprehensive picture of different categories of safety in the HRC context. [66] proposes an automated safety configuration for HRC regarding resources, processes, products, hazards, and the introduced safety behaviors model.

The distance between humans and cobots during work is a significant factor that impact human safety. As the distance between a human and a cobot diminishes, the cobot gradually slows down until it eventually comes to a stop. Several researchers addressed the safe distance between humans and cobots. [67] develops simulation software that finds a safe distance for a human operator and a cobot by GA. Two criteria regarding human safety are integrated into the HRC disassembly planning model proposed in [1]. These criteria are preventing unsafe task allocation to the human operator and ensuring a safe distance between human and cobot operators. [68] balances an HRC assembly line by solving constraint integer programs (SCIP), focusing on maintaining a safe distance between a human and a cobot. This distance depends on several factors: human and cobot reaction times and the required distance for the cobot to stop. The approach aims to allocate tasks in a manner that ensures the distance between the human and the cobot is greater than the safe distance, leading to faster cobot operation time and shorter process time.

Ergonomic risk Ergonomic risks in the workplace are a serious concern within the manufacturing industry. Failure to address these risks may result in work-related musculoskeletal disorders (WMSD) and other irreparable damage to the human skeletal system over time. According to the Occupational Safety and Health Administration (OSHA)[2], several activities at the workplace may potentially pose ergonomic risks. Applying overly strong force, frequently or continuously doing repetitive actions, and working in an inappropriate position are examples of these activities. Various approaches have been proposed to integrate ergonomic considerations into assembly/disassembly planning. Furthermore, in recent years and with the increasing use of cobots in manufacturing, a growing body of research has emerged to analyze the ergonomic risks in HRC industrial processes. [69] develops an HRC assembly line balancing model aiming to optimize both operation time and ergonomic risks. An integer planning model is developed in [70] to optimize assembly lines with respect to ergonomic risks. [71] proposes a fuzzy-based assembly line-balancing model that involves ergonomic risks. The proposed model considers four types of ergonomic risk: twisting the wrist, lifting, twisting the hip, and squatting. [72] conducts an experiment comparing several key performance indicators (KPIs), such as ergonomic risk, in the collaboration of a human operator and a cobot with a scenario in which the human operator solely works. Results show that the HRC strategy effectively reduces ergonomic risks and physical stress compared to manual disassembly.

[2]https://www.osha.gov/ergonomics/identify-problems

Technical feasibility The feasibility of an operation is a major issue in a disassembly process. Operators sometimes require advanced tools, and demolition machines, or even cannot disassemble some parts due to lack of strength. These parts remain incompletely disassembled and are transferred outside the cell. Therefore, it is essential to consider task feasibility in disassembly planning. A few researchers addressed the technical feasibility of disassembly during the task allocation process. Alrufaifi et al. [74] have computed the feasibility and direction of contact and non-contact disassembly based on extracted information from CAD models. They also generated a weighted graph showing the contact type and feasibility under six axes. The MGFEM method introduced in [78] evaluates the disassembly feasibility of a product in a partially destructive process. First, it computes the failure characteristics of the parts based on experts' knowledge and the type of failure. Then, a value is calculated based on three significant factors concerning disassembly feasibility. If this value is higher than a threshold, partial destructive disassembly is not feasible. Otherwise, the algorithm moves forward and uses a hybrid approach based on fuzzy logic and entropy to optimize the process. Moreover, [75,76] consider feasibility based on the physical precedence of a product's components. To find the optimum sequence, GA considers feasibility in addition to cost and environmental impacts as the objectives.

The current literature mostly focuses on financial factors but lacks a comprehensive sustainable method that integrates all aspects of sustainability into the HRC disassembly planning process. In this research, we introduce a novel sustainable framework for cobotic disassembly planning, which is based on a fuzzy-RL model. This model considers different elements regarding all three aspects of sustainability (economic, environmental, and social). These elements include operation time, tool change, energy consumption, circularity, ergonomic risk, and human safety in addition to technical feasibility. The hybrid fuzzy-RL method leverages fuzzy logic to cope with the uncertain nature of data while using the RL model to overcome uncertainties in the decision-making process.

To express the novelty of our proposed model, we compare it to recent studies in Table 5.1, focusing on the comprehensiveness of disassembly features involved in the planning process. While these articles incorporate various factors into the planning process, they use a limited range of features for devising the models. For example, consumed energy is not taken into account in any of the HRC studies. On the other hand, this research presents an HRC disassembly planning model, capable of making decisions in real-time with respect to the associated sustainable objectives.

Table 5.1 A qualitative analysis of the proposed model with the recent literature

	HRC	Real-time planning	Difficulty	Time	Circularity	Energy	Safety	Tool change	Ergonomic risk	Feasibility	Human skill
[48]	×	×	×	√	×	×	×	√	×	×	×
[60]	×	×	×	×	×	√	×	√	×	×	×
[59]	×	×	×	√	×	√	×	√	×	×	×
[133]	×	×	×	√	×	×	×	√	×	×	×
[88]	√	×	×	√	×	×	×	×	×	×	×
[1]	√	×	×	√	×	×	√	√	×	×	×
[86]	√	×	√	√	×	×	×	√	×	×	×
[128]	√	×	√	√	√	×	×	√	×	×	×
[140]	√	√	√	√	√	×	×	×	×	×	√
Our model	√	√	√	√	√	√	√	√	√	√	√

5.4 Problem statement

This research aims to present a new data-driven model for planning the disassembly process in the HRC setting. Considering objectives for the process, this model aims to effectively allocate tasks to the human and cobot, leading to the improvement of the efficiency and sustainability of the process. We define these objectives according to the sustainable criteria explained in Section 1.2. Equation 5.1 expresses the general objective function, where f_p and g_q represent the objectives to be minimized and maximized, respectively. In the following, each objective is explained in more detail.

$$\text{Minimize} \left(\sum_{p=1}^{P} f_p - \sum_{q=1}^{Q} g_q \right) \tag{5.1}$$

Task operation time is one of the most important objectives in a disassembly planning method. High profits are a result of reducing the total time of the process, depending on task execution time. Equation 5.2 denotes this objective, where T_{op} is operation times. The terms i and t_i refer to task i and its operation time, respectively. Moreover, N is the total number of tasks.

$$f_1 = T_{op} = \sum_{i=1}^{N} t_i \tag{5.2}$$

Additionally, a human operator may need to change tools based on the task requirements. Tool changing is a time-consuming action, affecting the total time of the process. Reducing the frequency of tool changes will decrease the overall processing time. Therefore, we consider minimizing the number of tool changes (T_{change}) as an objective in the planning model, as shown in Equation 5.3. Herein, $tool_i$ is the required tool for performing task i. δ is also a binary function that is equal to 1 if and only if the tools used for task i and task i+1 are different, and 0 otherwise. In this way, the model aims to determine the task sequence in such a way as to meet not only other objectives but also minimize the frequency of tool changes.

$$f_2 = T_{change} = \sum_{i=1}^{N-1} \delta \left(tool_i, tool_{i+1} \right) \tag{5.3}$$

Due to their physically demanding nature, disassembly processes often involve tasks that pose significant ergonomic risks (E_{risk}). These tasks have the potential to impact human operators' skeletal-muscular health in the long run. By reducing these risks, the potential for such health issues and human fatigue decreases, leading to enhanced performance. Hence,

an objective of the planning model is to minimize these risks, as indicated in Equation 5.4, wherein r_i is the ergonomic risk related to task i.

$$f_3 = E_{risk} = \sum_{i=1}^{N} r_i \tag{5.4}$$

With the increasing importance of the energy issue in the manufacturing industry, the amount of energy consumed by cobots has become more challenging. Although involving cobots in these processes has the potential to reduce costs and increase profitability, it is necessary to note that the inefficient use of cobots can lead to considerable environmental consequences. In other words, the more cobots are used, the greater their financial cost and energy footprint will be. As a result, we include minimizing energy consumption (E_{cons}) as an objective in the planning model, as shown in Equation 5.5, where e_i represents the consumed energy for completing task i.

$$f_4 = E_{cons} = \sum_{i=1}^{N} e_i \tag{5.5}$$

The level of circularity in the output plays an important role in addressing environmental concerns. Enhancing the quality of recovered parts (Q_{rec}) reduces the need for new parts in the remanufacturing processes. Consequently, it significantly decreases the consumption of raw materials. Therefore, improving the quality of the parts is a crucial primary objective of the planning process. As shown in Equation 5.6, one objective of the model is to maximize this quality. Herein, q_i presents the quality of recovered part i.

$$g_1 = Q_{rec} = \sum_{i=1}^{N} q_i \tag{5.6}$$

Ensuring the safety of human operators during disassembly processes is crucial because of potential exposure to hazardous tasks. Enhancing human safety increases operators' satisfaction and motivation, thereby boosting their performance. Hence, maximizing human safety is another objective of the planning model presented in Equation 5.7, wherein s_i represents the safety level associated with task i. Accordingly, given the hazardous nature of the tasks, the model aims to generate a task sequence that maximizes human safety.

$$g_2 = S_{human} = \sum_{i=1}^{N} s_i \tag{5.7}$$

HRC disassembly processes encompass many uncertain factors. One example is the vari-

able performance of human operators in executing tasks due to fatigue and distractions. Additionally, other uncertain factors such as machine failures, tool wear and tear, and any external disruptions may lead to variations in a task completion time. Given these scenarios, real-time decision-making based on current conditions is essential for task allocation, as any predetermined task sequence may become infeasible due to these uncertainties. Furthermore, the unstable conditions of EoL products are another uncertain factor, resulting from various elements such as damage, tear, and wear of parts due to different usage patterns. Developing an effective and efficient HRC disassembly planning model requires addressing these uncertain factors.

It is impractical to plan disassembly processes without considering the associated constraints of the problem. Equations 5.8-5.11 define these constraints. Equation 5.8 explains the limitation on the number of operations at a time, ensuring that each operator j can only perform one task i at a time. Here, $x_{i,j}$ is a binary variable that equals 1 if task i is assigned to operator j, and 0 otherwise. This constraint presents a practical limitation that an operator cannot perform multiple tasks at the same time as it is practically impossible. Equation 5.9 indicates that each operator can only use one type of tool per task, adhering to the assumption that working on each task requires only one tool. This assumption simplifies the process and reduces the complexity associated with changing tools in the model. Here, $tool_{i,t}$ is a binary variable that equals to 1 if performing task i requires tool t, and 0 otherwise. Equation 5.10 illustrates the constraint related to task dependencies on each other, in which task i precedes task k. S_i and S_k identify the start times of tasks i and k, respectively. Moreover, D_i represents the duration of task i. This equation ensures that task k should begin only after the completion of task i. This constraint effectively addresses the task sequence requirement and respects the precedence relationships between tasks in the planning process. Equation 5.11 ensures that a completed task i by an operator j cannot be reassigned, in which $y_{i,j}$ is a binary variable indicating whether a task i has been completed by an operator j or not. This constraint prevents disruptions from re-allocation completed tasks in the disassembly planning process, efficiently contributing to the resource management and time savings. Notably, this research assumes that each task is assigned to only one operator.

$$\sum_{i \in Tasks} x_{i,j} \leq 1, \quad \forall j \in Operators \tag{5.8}$$

$$\sum_{t \in Tools} tool_{i,t} \cdot x_{i,j} \leq 1, \quad \forall i \in Tasks, \quad \forall j \in Operators \tag{5.9}$$

$$S_i + D_i \leq S_k, \quad \forall i, k \in \text{Tasks where i precedes k} \tag{5.10}$$

$$\sum_{j \in \text{Operators}} y_{i,j} = 1, \quad \forall i \in \text{Tasks} \tag{5.11}$$

As explained in this section, the disassembly planning problem comprises multiple objectives, uncertainties, and constraints. In the following section, we discuss the developed approach to effectively address these challenges.

5.5 Methodology

This section presents our proposed sustainable HRC disassembly planning model. First, it introduces the employed graph-based modeling approach. Subsequently, the sustainable HRC disassembly method is explained. In the following, we discuss a sustainable multi-agent RL disassembly planning model and define its corresponding elements. Then, a fuzzy-based environment is introduced. Lastly, the entire architecture is comprehensively discussed.

Graph representation of products Representing a product architecture is a crucial step in disassembly planning. We use a graph-based representation approach introduced in [140] to model a product architecture. Here, each component is identified as a node, while each connection between two components is represented as an edge. This approach categorizes a product's components into different levels according to their distance from the origin node of the product, the initial point which all nodes and edges originate from. Figure 5.1 illustrates a product' graph, wherein nodes in the same level have the same color. In this case, a process starts by allocating components from the outermost level. After disassembling a component at each level, the algorithm continues by choosing a component at the same level or a component in the next level connected to the previously disassembled one. The algorithm shifts to the next level after completing all tasks in one level. This graph-based approach filters out certain components from all candidate components during the allocation process, yet it does not select the next component. Hence, we propose an RL model, discussed in Section 5.5.1, to achieve the optimum task sequence.

Interchanging the collaborative and inactive modes Occasionally, only highly difficult and complex tasks are available during disassembly processes. Performing such tasks with a cobot may negatively impact both its physical structure and the product. In these

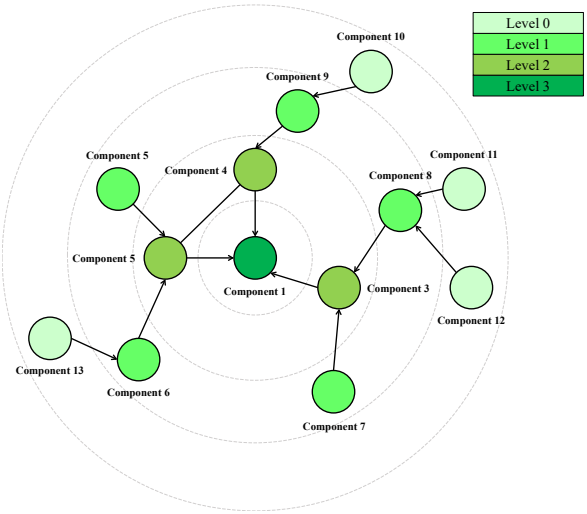


Figure 5.1 The graph representation approach

cases, the human operator should work exclusively, while the cobot should be stopped and wait until simpler tasks become available. As a result, human-robot interaction shifts from collaboration mode to inactive mode, and the multi-agent model transitions to a single-agent model. To address this issue in modeling, we adopt the approach presented in [140]. This approach selects tasks with the lowest difficulty and highest technical feasibility among the available ones. If the difficulty is categorized as low and the feasibility as high, there is at least one task that can be assigned to the cobot, allowing the process to use both the human and cobotic operators. Otherwise, the cobot stops, and the process switches to inactive mode. Section 5.5.2 provides further details on these categories.

Sustainability driven HRC disassembly planning This research proposes a sustainable development triangle in the context of HRC disassembly shown in Figure 5.2. Social, environmental, and cost impacts are three sides of this triangle. The social side includes human safety along with ergonomic risks. We consider working with hazardous substances as a harmful factor to the safety of human operators. Thus, the proposed model seeks to allocate tasks associated with hazardous materials to cobots. The environmental side comprises the circularity of parts and cobot energy consumption. The cost side consists of operation time and tool change. As shown in the center of the triangle in Figure 5.2, sustainable development in this area is impossible without considering the feasibility of tasks for both cobots and humans. A disassembly process should be planned to avoid assigning infeasible tasks to both human operators and cobots. Involving operators with these tasks results in prolonged execution times and poor profitability. Therefore, another objective of the planning model

is to optimize the process by avoiding the allocation of such tasks to the operators.

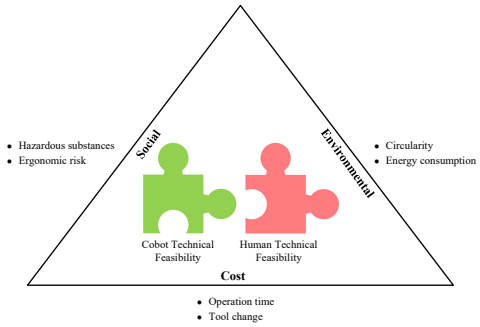


Figure 5.2 Sustainable development disassembly planning triangle

5.5.1 Proposed RL-based sustainable HRC disassembly planning model

This research proposes a sustainable HRC disassembly planning fuzzy-RL model. It comprises two agents, a human agent and cobot agent, responsible for allocating tasks to human and cobot operators. In this approach, the agents recursively interact with an environment. During an iteration, each agent is in a state s_t (time step t), it then selects an action a_t based on a policy π. Next, the agent receives a reward r_{t+1} (time step t+1) as a feedback signal from the environment in response to the executed action. After that, the agent proceeds to a new state s_{t+1} and repeats this cyclic process. This research study formulates each iteration by using Markov decision process (MDP), a tuple that includes four parts: s_t, a_t, r_{t+1}, and s_{t+1}. As this process evolves, the agent learns to choose an action in each state resulting in a higher reward value. In the following, the agent finally learns the optimum policy π^* that leads to taking the optimum action in each state, resulting in the maximum immediate and long-term reward values. The optimum policy formula is shown in Equation 5.12, in which the term T refers to the total time steps.

$$\pi^{\star} = \arg\max_{\pi} \mathrm{E}\left[\sum_{t=0}^{T} R_t\right] \tag{5.12}$$

Below, we explain different elements of the RL model tailored specifically to meet the requirements and objectives of this research study.

State In RL, the state vector represents the current condition of the problem. We determine the state vector based on features of the product components, which should be disassembled. These features are the difficulty of component disassembly, the feasibility of operations, the safety of operations concerning working with hazardous substances, and ergonomic risks.

We obtain the difficulty score of a component disassembly through the linear summation of various component characteristics. The elements of this linear combination are strength, weight, size, positioning, accessibility, destructive, liaisons scores, liaisons properties, and number of connections with other components. Within this context, strength addresses the needed force for an operation. The weight and size of a component are also two important factors in its disassembly. According to [82], positioning is defined as correctly situating a tool on the connection of a component before initiating the operation. Accessibility refers to the capability to reach a specific area by hand or instrument [80]. The difficulty of a task is also significantly impacted by whether it requires a destruction operation or not. Moreover, we assign values ranging from one to two to each connection's properties, while distinct scores are allocated to each connection type. Also, the number of connections of a component with other components is an impactful factor in its disassembly difficulty. The greater number of connections for a component makes its disassembly more challenging and complex. Difficulty scores are integrated into the states of both human and cobot agents. This research takes into account the feasibility of tasks for both human operators and cobots, noting that the feasibility values of tasks for cobots and humans are different. Accordingly, state vectors of human and cobot agents include the feasibility values. We also consider binary values related to hazardous substances in the state of the human agent. Each value corresponds to a task, and it represents whether the task involves dangerous materials or not. On the other side, cobots are capable of working on such tasks regardless of the included dangerous substances. Therefore, we do not incorporate the values concerning hazardous substances in the state vector of the cobot agent. As well, we consider the same procedure for ergonomic risks by defining related binary values in the state of the human agent. Similar to the hazardous substances, the ergonomic risk values are not included in the cobot agent's state. It is noteworthy that after completing each task, its corresponding values in states would be zero.

Action The action space covers all currently available tasks for allocating to human and cobot operators in each iteration. Accordingly, an action refers to the next task. It is notable that an agent should select a new action in each iteration. The agent cannot repeat the same action during a process, as each chosen action addresses a task that has already been completed.

Reward The reward functions are specifically tailored to match agents' objectives. As the goals and responsibilities of humans and cobots in the disassembly process are different, the associated reward functions are customized accordingly. Equations 5.13 and 5.14 define the

reward functions of the human and cobot agents, respectively. The term R_q addresses the recovered component quality, representing the circularity of parts. The time of the operation is also denoted by the penalty term R_t. Furthermore, the binary value R_H addresses whether the operation involves hazardous substances or not. The term R_E refers to the ergonomic risk associated with the operation as well. Ergonomic risks are categorized into three levels of 'Low', 'Medium', and 'High', corresponding to different tasks. In our model, we assigned numerical values of 0, 0.5, and 1 to these categories, respectively. The detailed assessment of these risks is beyond the scope of this research, and we addressed it in a separate study [141]. In the current article, we use the ergonomic risk values in the utilised dataset and integrated them into the proposed model. The penalized reward R_{Energy} refers to consumed energy for the corresponding executed operation. We assume that with each operation performed by the cobot, the term R_{Energy} increases by one unit. Since our model was evaluated in a simulated environment, we are unable to measure actual energy consumption. However, this assumption allows us to demonstrate the concept effectively. The penalty phrase R_{Tool} addresses tool change and is a binary value that represents whether the operation requires a tool change or not. Lastly, the reward R_p is a penalty term to incentivize the human agent to choose more complex tasks, while this term also motivates the cobot agent not to select challenging tasks.

The impact of each term on the reward function is identified by its coefficient. We can modify the importance of each objective in the decision-making process by adjusting its respective coefficient. For instance, by increasing the coefficient of a particular term, the model puts more attention to selecting actions that align with maximizing the reward associated with that term. Moreover, the coefficients serve to unify the objectives prior to aggregation, as they have different units. This study assigns a value of one to all these coefficients. However, they can be changed according to the conditions and goals of the problem.

$$R_{human} = \alpha_q \cdot R_q - \alpha_t \cdot R_t - \alpha_H \cdot R_H - \alpha_E \cdot R_E - \alpha_{Tool} \cdot R_{Tool} - \alpha_p \cdot R_p \tag{5.13}$$

$$R_{cobot} = \alpha_q \cdot R_q - \alpha_t \cdot R_t - \alpha_{Energy} \cdot R_{Energy} - \alpha_p \cdot R_p \tag{5.14}$$

This study aims to present an RL model for sustainable HRC disassembly planning. In this context, sustainable decision-making refers to the ability to manage and trade-off between objectives in task allocation, arranging the sequence of tasks in such a way that all objectives are optimally considered. Here, the reward functions are the combination of terms corresponding to the sustainability pillars, and the reward values represent the models' per-

formances in task allocation with respect to the sustainability objectives. Given the recursive nature of RL, this strategy drives each model to make more sustainable decisions to achieve higher reward values. Since the proposed framework comprises two agents with different reward functions, sustainability results from maximizing the reward values of both agents. Hence, the convergence of the reward values of both agents to the maximum values at the same time indicates the good performance of the framework in the sustainable task allocation process.

Deep Q-network We develop the multi-agent planning architecture by using two deep Q-networks (DQN), representing the human and cobot agents. DQN is a model-free RL neural network, working based on the Q-learning theory. DQN approximates Q-values for each state-action pair to select an action with the highest Q-value based on a policy π. It targets to achieve an optimum Q-function $\mathrm{Q}^*(\mathrm{s}, \mathrm{a})$, given by Equation 5.15, to choose the most effective action in each state. It results in the maximum cumulative immediate and long-term reward values. The term γ serves as the discount factor of the equation, balancing the importance of short-term and long-term reward values.

$$\mathrm{Q}^*(\mathrm{s}, \mathrm{a}) = \max_{\pi} \mathbb{E}\left[\mathrm{R_t} + \gamma \mathrm{R_{t+1}} + \gamma^2 \mathrm{R_{t+2}} + \ldots \mid \mathrm{s_t} = \mathrm{s}, \mathrm{a_t} = \mathrm{a}, \pi\right] \tag{5.15}$$

DQN contains two sub-neural networks: prediction and target networks. While the prediction network approximates the Q-values for the current state s, the target network estimates the Q-values regarding the next states, shown by $\mathrm{Q}(\mathrm{s}', \mathrm{a}')$. The Q-values are then continuously updated through Equation 5.16, where the term α refers to the learning rate of the process. The associated parameters of the model are summarized in Table 5.2.

$$\mathrm{Q}(\mathrm{s}, \mathrm{a}) \leftarrow \mathrm{Q}(\mathrm{s}, \mathrm{a}) + \alpha[\,\mathrm{R}(\mathrm{s}, \mathrm{a}) + \gamma \arg\max_{\mathrm{a}'} \mathrm{Q}(\mathrm{s}', \mathrm{a}') - \mathrm{Q}(\mathrm{s}, \mathrm{a})\,] \tag{5.16}$$

Equation 5.17 presents the loss function of the process, in which θ and θ^- denote the parameters in the prediction and target networks, respectively. Stochastic gradient descent is typically adopted to update the parameters of the prediction network instead of calculating the complete summation. At intervals of $\mathrm{T_{target}}$ time, the target network parameters are updated based on the prediction network parameters. Furthermore, this research adopts the ϵ-greedy algorithm to balance the exploitation-exploration dilemma in decision-making (action selection). In exploitation mode, actions are selected based on previously acquired knowledge, whereas they are chosen randomly in exploration mode. This way, we determine a value for a parameter ϵ at the beginning of the process. At each time step, if the ϵ value

is higher than a random value obtained from a Gaussian distribution, an action is randomly selected (exploration). Otherwise, the RL agent selects an action according to the proposed approach (exploitation). At the end of each episode, the ϵ value decreases by a discount factor, reducing the exploring frequency as the learning process progresses. Therefore, the agent gradually relies on its knowledge rather than exploration to select actions during the evolution of the learning process.

$$\mathrm{L}(\theta) = \sum_{\mathrm{s,a,r,s'}} \left(\mathrm{r} + \gamma \max_{\mathrm{a'}} \mathrm{Q}\left(\mathrm{s'}, \mathrm{a'}; \theta^{-} \right) - \mathrm{Q}(\ \mathrm{s}, \mathrm{a}; \theta) \right)^{2} \tag{5.17}$$

5.5.2 Fuzzy-based disassembly environment

A disassembly process is associated with uncertainties that may arise from several factors. Among these factors are the long-term use of the product, environmental conditions, technical proficiency of operators, as well as tools and methods employed in the process [142]. The uncertainties may be exhibited in both the decision-making process and data. This research tackles the uncertainty in making decisions by developing a real-time RL-based planning model. In terms of data, the mentioned factors result in the existence of uncertainty in parameters, such as the quality of recovered parts. Modeling these parameters is essential as they should be considered in the decision-making process. These parameters express subjective concepts, and consequently, it is difficult to numerically represent them due to the lack of prior knowledge about them. It is also impractical to determine exact values for these parameters because of their uncertain nature.

This research uses a fuzzy logic method to effectively model uncertain parameters of the problem (difficulty, feasibility, operation time, and recovered quality). Uncertain parameters, along with their associated probability density functions, can be approximated using historical data. However, this method proves inefficient for this problem due to the lack of historical data on product disassembly planning issues. On the other hand, fuzzy-based approaches can model parameters under uncertainty in a sensory and experiential manner based on expert knowledge or experience-based estimates, even in scenarios with insufficient historical data. Here, we design a fuzzy-based environment to compute reward values in response to an executed action. It is notable that the reward functions consists of fuzzy and crisp parameters. Difficulty, feasibility, operation time, and recovered quality are fuzzy parameters, while consumed energy, tool change, safety, and ergonomic risks are crisp parameters. We compute the operation time and recovered quality crisp values with a fuzzy-inference system (FIS) according to difficulty and feasibility values. Then, we sum all of the crisp values to

Table 5.2 Parameters in the proposed fuzzy-RL model

Parameter	Description
π	Policy in Equation 7.8
γ	Discount factor in Equation 7.9
R_q	The reward term related to the quality of the recovered component
R_t	The reward term representing the operation time
R_H	The binary reward value presenting whether a task involves hazardous substances or not.
R_{Tool}	The binary reward value corresponds to whether this operation requires a tool change or not.
R_E	The reward term corresponding to ergonomic risk of the operation
R_P	The reward term aimed at penalizing the agents
R_{Energy}	The reward term presenting the consumed energy for the executed action by the cobot
M_h	The human agent replay memory
M_c	The cobot agent replay memory
α	Learning rate in Equation 7.9
ϵ	ϵ-greedy algorithm parameter
s_t^h, a_t^h, r_t^h	The corresponding state, action, and reward of the human agent at time t
s_t^c, a_t^c, r_t^c	The corresponding state, action, and reward of the cobot agent at time t
T_{target}	Number of iterations required for updating the target network parameters

calculate the reward scores. As shown in Figure 5.3, the fuzzy-based environment includes three steps. It first executes the action selected by an RL agent. Subsequently, the FIS fuzzifies the uncertain parameters consequencing to the executed action. The FIS infers the data by using pre-defined fuzzy rules to compute outputs, which are then defuzzified. Finally, the defuzzified parameters are combined with the non-fuzzy parameters according to Equations 5.13 and 5.14 to form the reward values. After that, the RL model receives the reward value from the environment.

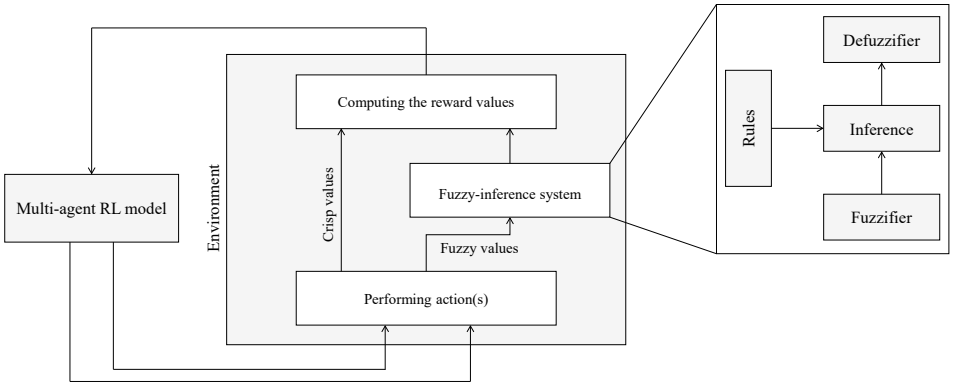


Figure 5.3 The proposed fuzzy environment

This study models uncertain parameters with fuzzy triangle numbers due to their suitability for representing the inherent imprecision in factors related to disassembly sequence problem. Accordingly, it considers three distinct membership functions - low, medium, and high -

for each parameter. Figure 5.11 in the Appendix illustrates the triangular fuzzy numbers, including difficulty, feasibility, operation time, and quality of recovered components. The bounds of these fuzzy numbers are properly determined according to experts knowledge regarding the problem and the data used to evaluate the model.

This research categorizes human operators' skill levels into three groups: Acceptable, Medium, and High. These groups are defined based on the following characteristics:

- Acceptable: Novice operators possess basic skills in disassembly tasks and have recently begun working on disassembly processes. These operators typically complete tasks in a longer time and have a higher possibility of making errors.
- Medium: Operators at this skill level demonstrate higher proficiency than novice operators, with the ability to successfully complete more difficult tasks.
- High: Proven experts with extensive experience in performing challenging and difficult tasks. These operators generally achieve fewer errors, higher performance, and faster operation times.

This classification demonstrates the varying performance of humans in executing disassembly tasks. Accordingly, the environment provides different feedback based on human operator's skill level. This assumption allows us to effectively provide a more comprehensive simulation and validation regarding the proposed planning model. Here, we consider three scenarios in which human operators with different skill levels collaborate with a cobot in a process. As a result, the fuzzy environment provides three types of rules for human operators based on their skill levels in addition to rules for cobots. Algorithm 5 in the Appendix explains the fuzzy rules related to environment modeling for a cobot and different skilled human operators. We define these rules according to experts' knowledge regarding the conditions of this problem.

5.5.3 The proposed model architecture

This section presents a comprehensive overview of the proposed architecture, in which the human and cobot agents interact with a fuzzy-based environment. The block diagram provided in Figure 5.4 explains this interaction in more detail. In the initial phase, we set up the agents with randomly chosen weights and biases. In the following, the agents first choose actions (tasks) a_t^h and a_t^c in states s_t^h and s_t^c, respectively. Each of the agents receives a reward value from the fuzzy-based environment whenever the corresponding task is completed. Here, we show the completion time regarding human and cobot tasks by $t^{'}$ and $t^{''}$. After

that, the respective state shifts to a new state, shown by $s^h_{t'}$ and $s^c_{t''}$ for the human and cobot agents in order. This process continues by storing an MDP corresponding to each iteration in the associated replay memory, denoted as M_h and M_c. For each agent, a batch of data is sampled from the associated replay memory to update weights and biases of the corresponding agent. While an operator is executing a task, the other operator may have finished their respective task. In this problem, we do not want to reallocate executed tasks to operators. Therefore, the state of an agent should be updated according to possible changes caused by another agent (zero values are assigned to the indices corresponding to the task executed by one agent in the state vector of other agent too). Then, each agent chooses a new action, and the algorithm moves to the next time step. Subsequently, the process will continue until all tasks are completed. Notably, this algorithm allocates tasks in real-time based on online conditions of the process rather than generating a hypothetical pre-determined task sequence. Therefore, the model is capable of dealing with uncertain situations that may change a process plan. Since this model works in a dynamic and adaptive form, it can adjust to potential variations in a non-static environment.

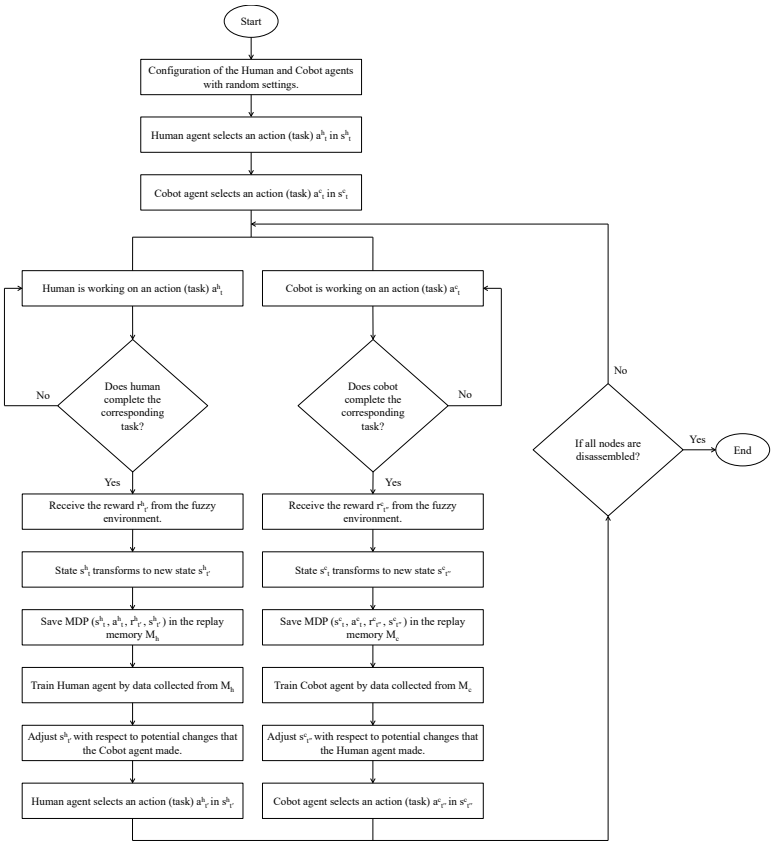


Figure 5.4 Architecture of the sustainable HRC disassembly framework using fuzzy-RL

5.6 Experiments

This section discusses the experimental setup and results. First, we explain a case study, a metric to evaluate the proposed model, and parameter settings. Then, we discuss the results of conducting three scenarios and different sensitivity analyses concerning the three most important parameters of the model. These experiments focus on evaluating the strength and adaptability of the model to a variety of simulated and application conditions. We also perform a sensitivity analysis under uncertain conditions to assess the model's capacity for dynamic decision-making. Finally, we present a comparative analysis with respect to a baseline model.

5.6.1 Experimental setup

Data To evaluate the proposed approach, we use the hard disk drive data provided in [1]. Figure 5.5 illustrates the graph of the product and the geometric relationships among its components. Moreover, Table 7.7 illustrates the features considered for components of this product. Each row represents a component, while each column corresponds a specific feature.

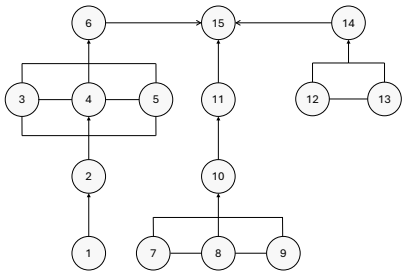


Figure 5.5 The structure of the product introduced in [1]

Evaluation protocol This research study measures the model's performance in different scenarios through the cumulative reward value of each episode. As shown in Equation 5.18, this value is achieved by summing the reward value of each time step during an episode, which contains T number of time steps. Gaining more reward values implies that the agents effectively learned to make decisions according to the objectives included in the reward functions.

$$\mathrm{R} = \sum_{\mathrm{t}=1}^{\mathrm{T}} \mathrm{r}_{\mathrm{t}} \tag{5.18}$$

Table 5.3 The utilised case study and its corresponding features values

Component number	Positioning	Accessibility	Weight	Strength	Size	Destructive	Liaisons scores	Liaisons properties	Hazardous substances	Ergonomic risk	Human feasibility	Cobot feasibility	Tool
1	5	1	2.4	4	3.5	1	10	2	0	Low	19	0	1
2	1.2	1.6	2	1	2	0	2	1.2	0	Low	19	0	1
3	1.2	1	2	2	2	0	3	1.3	0	Low	0	0	2
4	2	2	2.2	1	4	0	2	1.4	0	Low	0	0	1
5	1.2	1.6	2.2	2	3.5	0	4	1	0	Low	19	0	3
6	5	1	2	1	2	0	2	1.1	1	Low	19	19	1
7	2	1	2.4	4	3.5	1	6	1.4	0	Low	19	0	1
8	1.2	2	2	2	2	0	3	1.2	0	Low	19	19	1
9	2	1.6	2.2	2	2	0	2	1.1	0	Low	19	0	3
10	2	1	2	1	2	0	4	1.5	0	Low	19	19	2
11	1.2	1.6	2.2	1	2	0	2	1.7	0	Low	19	0	1
12	1.2	1	4	4	4	1	10	1.8	0	Low	19	0	1
13	1.2	2	2	1	2	0	2	1.1	0	High	19	19	3
14	1.2	2	2.2	2	2	0	3	1.9	0	High	19	19	1
15	5	2	4	2	4	0	4	1.5	0	Low	19	19	2

Parameter settings This research uses two feedforward layers with eight neurons for employed neural networks. It considers the parameter α value of 0.2, while the T_{target} value is set to 50. Furthermore, a batch size of 50 is utilised to update the weights and biases of DQN. This research also conducts sensitivity analysis to efficiently select the most optimum values for ϵ, γ, and replay memory size parameters, as they have the greatest impact on the proposed framework's performance.

5.6.2 Results and discussion

As this study categorizes human operators into three groups based on their expertise, it defines three human agents, which represent the roles of these groups in the RL scheme. We also design three scenarios involving the collaboration of a cobot with different types of human operators.

We analyze the model's performance by examining several sensitive parameters that have the most impact on the model's output. The sensitivity analysis is based on the parameters ϵ, discount factor γ, and replay memory size. Generally, RL models work according to the exploration-exploitation dilemma, a trade-off between choosing the best decision based on previous knowledge or searching to find better options. In the ϵ-greedy approach, the parameter ϵ sets this trade-off by determining the exploit and search iterations. This is a sensitive parameter as poor exploring may heavily fluctuate the reward values. Furthermore, focusing solely on exploiting may lead to a long learning process. The compromise between immediate and long-term reward values in decision-making is crucial. In this case, the parameter γ determines this sensitive trade-off. Additionally, the performance of the model may be affected based on the size of the replay memory as it dynamically samples data to train the sub-networks. Selecting big or small sizes for the replay memory may decrease the sampling performance as well as the learning process quality.

We prioritize the analysis of these parameters according to their importance. Hence, we examine the parameters ϵ, γ, and memory size in that order. It is worth noting that when changing any of these parameters, the remaining ones are held at their default configurations intact.

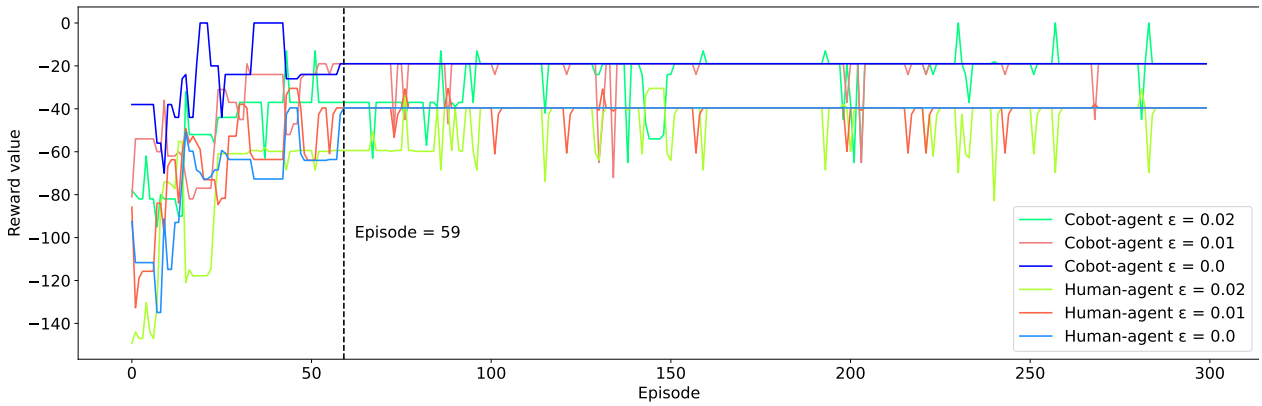
Scenario I: an acceptable-skill-level human operator Figure 5.6 illustrates the reward values for various parameter settings in this scenario. Sub-figure 5.6a shows the reward plots for different values of the parameter ϵ. The increase in the values, leading to further explorations by the agents, results in the appearance of more fluctuations in the reward plots. Consequently, it prevents the convergence of reward plots for both agents. In this

case, searching in the feature space not only degrade the performance of the model but also prevents its convergence. Therefore, the most optimum value for the parameter ϵ is 0, which the model does not explore. Sub-figure 5.6b illustrates the reward values based on different γ parameter configurations. The reward plots indicate slower convergence for γ parameter values of 0.01 and 0.1 compared to 0.001. However, the final convergence value in these cases is greater than when the γ parameter is set to 0.001. To select the optimal replay memory size, we effectively analyze the model's performance with three size settings: 50, 100, and 200. As shown in sub-figure 5.6c, the model converges faster by setting the size to 200. Moreover, the reward plots exhibit poor convergence when employing the size of 50. With this setting, the model experiences insufficient sampling quality, leading to a less effective learning process.

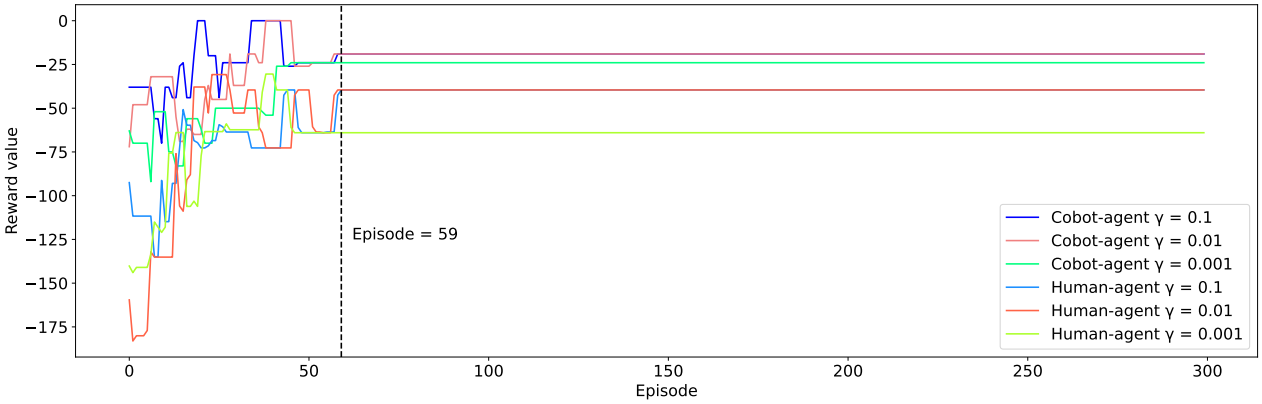
Scenario II: a medium-skill-level human operator Similar to the previous scenario, we obtain reward values with considerable fluctuation when the agents work in the exploration mode. Hence, the most optimum result, as shown in Sub-figure 5.7a, is achieved when the parameter ϵ value is 0. Furthermore, the reward plots in sub-figure 5.7b depict the model's performance across the variation in the parameter γ. The presented results reveal that when the values are reduced, which decreases the long-term reward impact in the decision-making process, the model's performance also declines. In terms of replay memory size, the plot shown in sub-figure 5.7c shows that the model with the size of 100 exhibits much faster convergence compared to models with the sizes of 50 and 200.

Scenario III: a high-skill-level human operator Sub-figure 5.8a illustrates the reward plots across various values of the parameter ϵ. Similar to the findings in Scenarios I and II, the agents achieve best performance when only engaging in exploitation. The reward plots depicted in sub-figure 5.8b are obtained by varying the parameter γ with three different values. By setting the value to 0.1 or 0.01, the plots converge more quickly. As demonstrated in sub-figure 5.8c, the agents meet faster convergence when the replay memory size is set to 100, in line with the results observed in Scenario II.

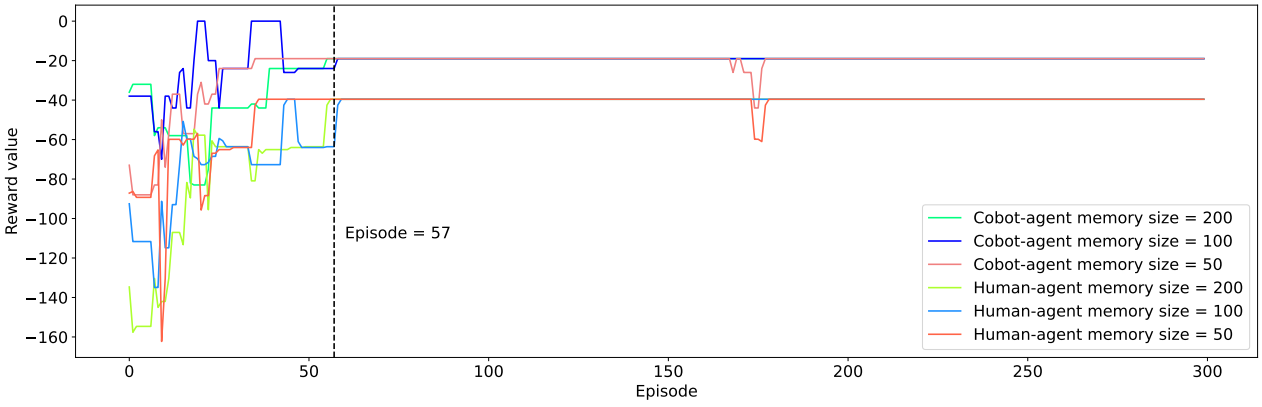
Sensitivity analysis of the parameter ϵ The exploration's purpose is to enhance the model's ability in decision-making by guiding it away from local optima, mitigating overfitting resulting from decisions made with limited knowledge, and introducing new observations to improve its knowledge. The results in Figures 5.6a-5.8a show that increasing the ϵ values leads to more fluctuations in the reward plots, resulting in slower convergence. In each scenario, the most stable result is achieved with $\epsilon = 0$, where the agents do not explore.



(a) Parameter ϵ sensitivity analysis

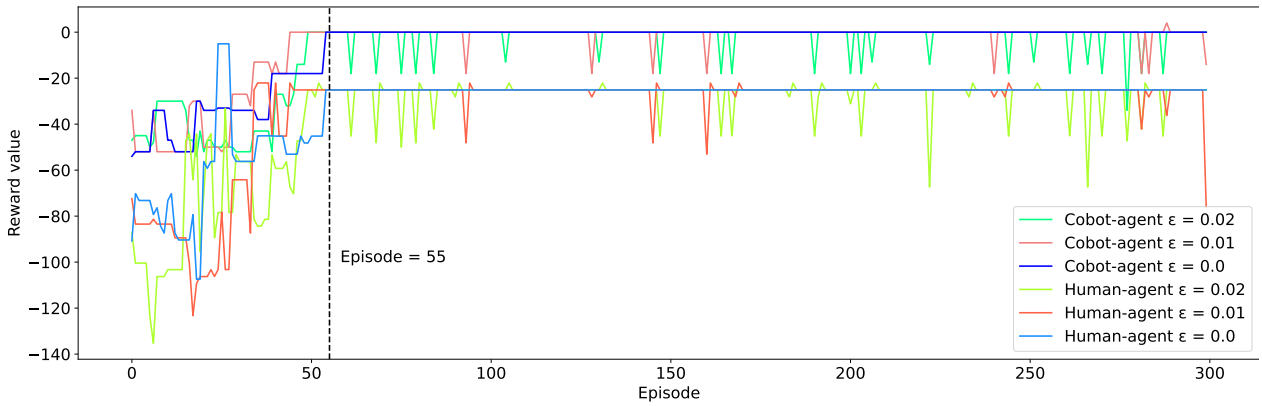


(b) Parameter γ sensitivity analysis

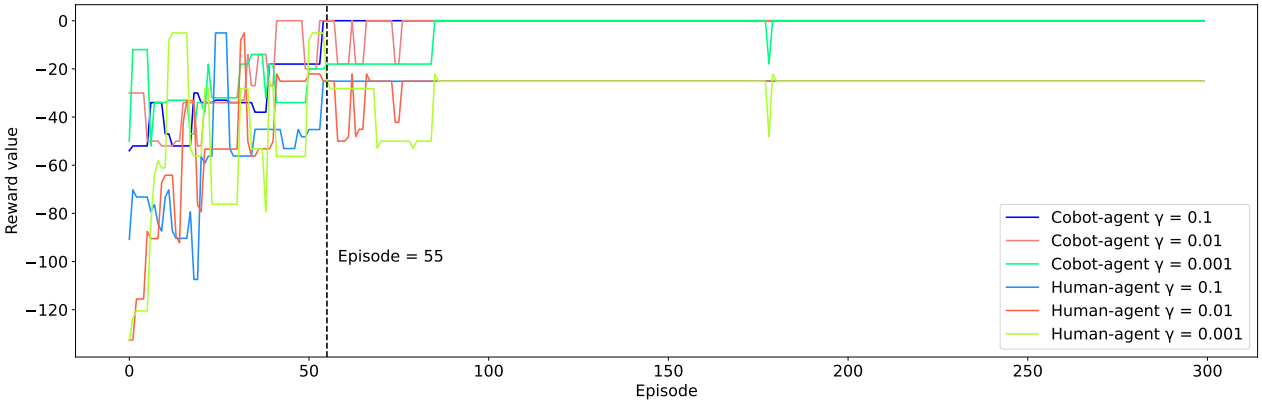


(c) Replay memory size sensitivity analysis

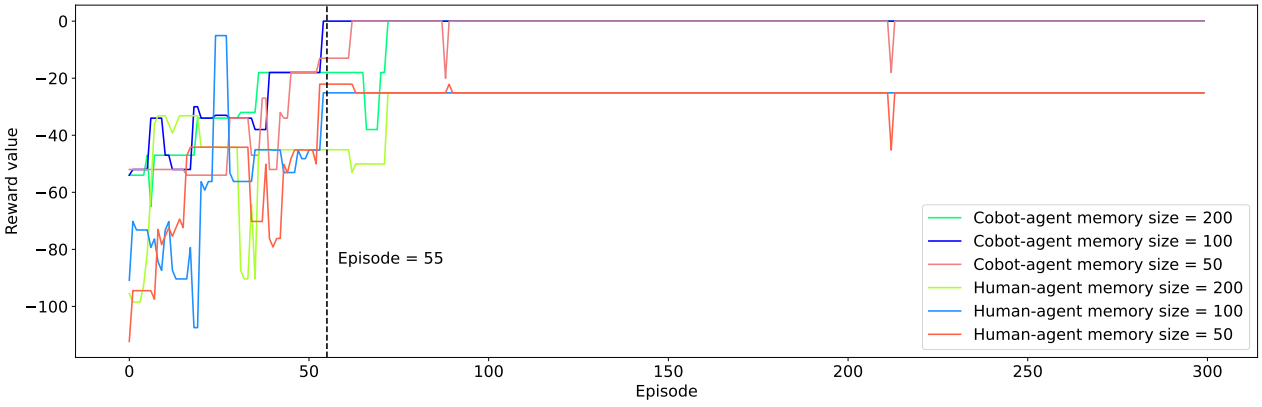
Figure 5.6 Sensitivity analysis in Scenario I



(a) Parameter ϵ sensitivity analysis

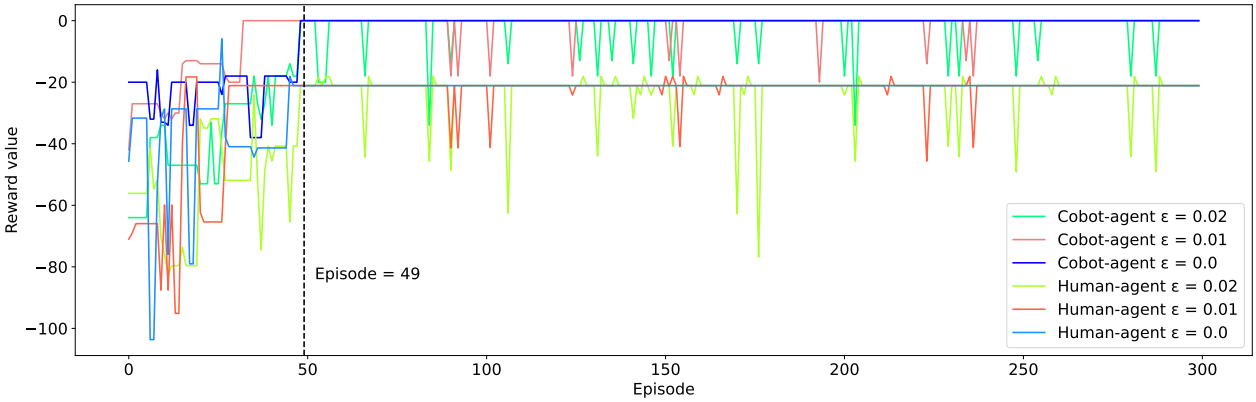


(b) Parameter γ sensitivity analysis

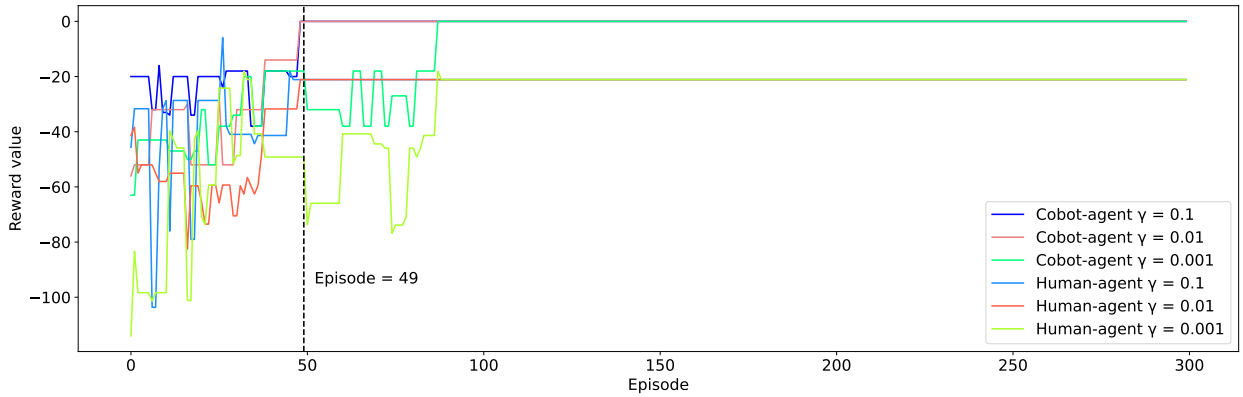


(c) Replay memory size sensitivity analysis

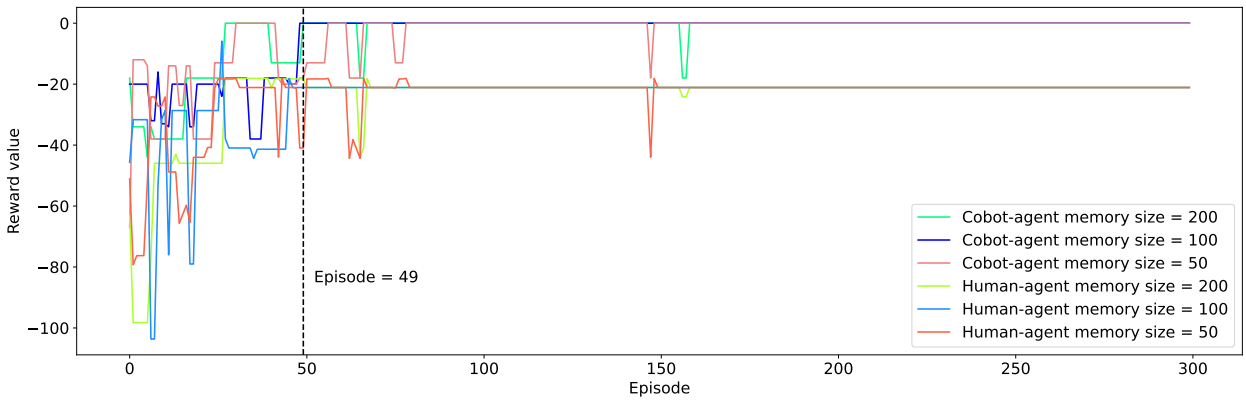
Figure 5.7 Sensitivity analysis in Scenario II



(a) Parameter ϵ sensitivity analysis



(b) Parameter γ sensitivity analysis



(c) Replay memory size sensitivity analysis

Figure 5.8 Sensitivity analysis in Scenario III

The results demonstrate that the model does not face the challenges of interacting with the environment that exploration can address. The small size of the problem space is a key factor contributing to the model's stable performance without requiring exploration. As the problem space expands, the model potentially encounters the mentioned obstacles, and exploring may improve its learning performance in such cases.

Sensitivity analysis of the parameter γ By decreasing the γ values, the stability of the reward plots decreases, as shown in Figures 5.6b-5.8b. In this case, the agents rely more on the immediate reward values in decision-making, with the consideration for long-term rewards. This setting results in poor performance of the model across scenarios, highlighting the importance of accounting for long-term rewards in planning. Accordingly, due to environmental characteristics, emphasizing immediate rewards alone diminishes the model's planning performance in this context. Moreover, due to the robustness of the model for $\gamma = 0.1$, this setting may result in broader generalizability across various environments and scenarios in this problem domain.

Sensitivity analysis of the replay memory size As depicted in Figure 5.6c, the model achieves its highest performance with a memory size of 200 in Scenario I. However, the model attains its best performance using a memory size of 100 in Scenarios II and III, as illustrated in Figures 5.7c and 5.8c. Additionally, our investigation confirms that a size of 50, resulting in a reduced number of observations in the memory, inhibits the learning process. Optimal replay memory provides sufficient observations for the model, enhancing the sampling process and learning performance. Given the negligible difference (2 episodes) between memory sizes of 100 and 200 in Scenario I, it can be concluded that a memory size of 100 exhibits greater generalizability compared to sizes of 50 and 200 in this problem domain. On the other hand, while the model meets its highest performance in Scenario I with a size of 200, this configuration does not yield efficient results in Scenarios II and III. An inefficient increased size results in greater variance in the data, leading to instability in the learning process. Hence, a size of 200 exhibits low generalizability in this context.

Sensitivity analysis on decision-making under uncertain conditions This analysis aims to validate the ability of the model to make decisions in uncertain conditions. As shown in Equation 5.19, we represent the uncertain environment by considering the variable time required to complete each task. In this context, T_{base} denotes the fixed time needed to complete a task, and T_{var} represents the additional variable time, caused by uncertain factors. The total time to complete the task, T', is calculated as the sum of the fixed

and variable times. T_{var} is selected from a 'Variations' vector based on a uniform discrete probability distribution, as shown in Equations 5.20 and 5.21. This vector has a length of 50 and contains the numbers 1, 2, 3, 4, and 46 instances of 0. Moreover, N represents the length of the vector, and δ is a binary function. As shown in Equation 5.22, δ becomes 1 if T_{var_j} equals x_i, and 0 otherwise. Consequently, the probability of sampling the numbers 1, 2, 3, or, 4 is 0.02, while the probability of sampling the number 0 is 0.92.

$$T^{'} = T_{base} + T_{var} \tag{5.19}$$

$$\text{Variations} = [T_{var_0}, T_{var_1}, \ldots, T_{var_N}] = [\underbrace{0, 0, \ldots, 0}_{\text{46 numbers}}, 1, 2, 3, 4] \tag{5.20}$$

$$T_{var} = P\left(X = x_i\right) = \frac{1}{N}\sum_{j=1}^{N} \delta\left(x_i, T_{var_j}\right) \tag{5.21}$$

$$\delta\left(x_i, T_{var_j}\right) = \begin{cases} 1 & x_i = T_{var_j} \\ 0 & x_i \neq T_{var_j} \end{cases} \tag{5.22}$$

We implemented the model developed in Scenario I within an environment comprising the mentioned uncertainties. Figure 5.9 illustrates the total rewards for both agents in this sensitivity analysis. Despite minor fluctuations, the rewards have achieved a relatively stable level, indicating the model's strong performance in decision-making, even with variable task times.

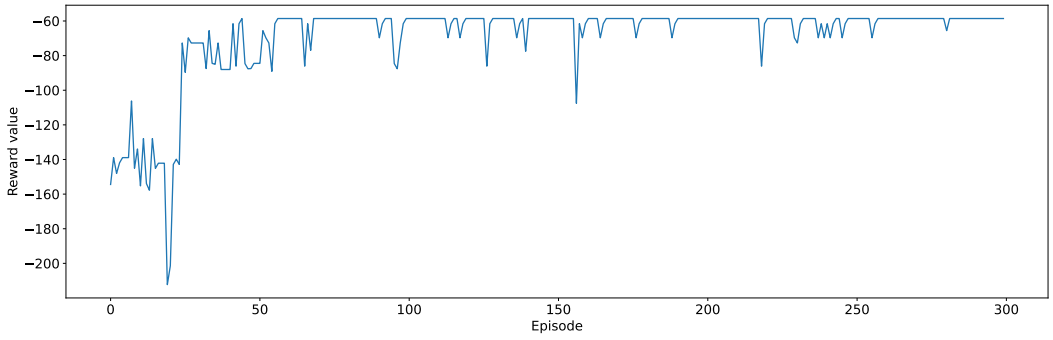


Figure 5.9 The reward values for sensitivity analysis on decision-making under uncertain conditions

Comparative analysis This part evaluates the proposed model by comparing it with the model presented in [1] as we utilise a case study from the referenced article. We use convergence time as a quantitative evaluation metric. The purpose of this comparison is not to benchmark execution time or computational efficiency under identical conditions, but rather to qualitatively assess whether the proposed RL-based approach can achieve bounded convergence behavior comparable to a classical method in the same problem context. Figure 5.10 compares the convergence times of our model under three scenarios with the best convergence time of the baseline model. The baseline model converges in 120 iterations with the optimum setting, whereas our model achieves convergence in 57, 55, and 49 iterations, respectively. Therefore, the developed model in this research converges much faster compared to the baseline paper.

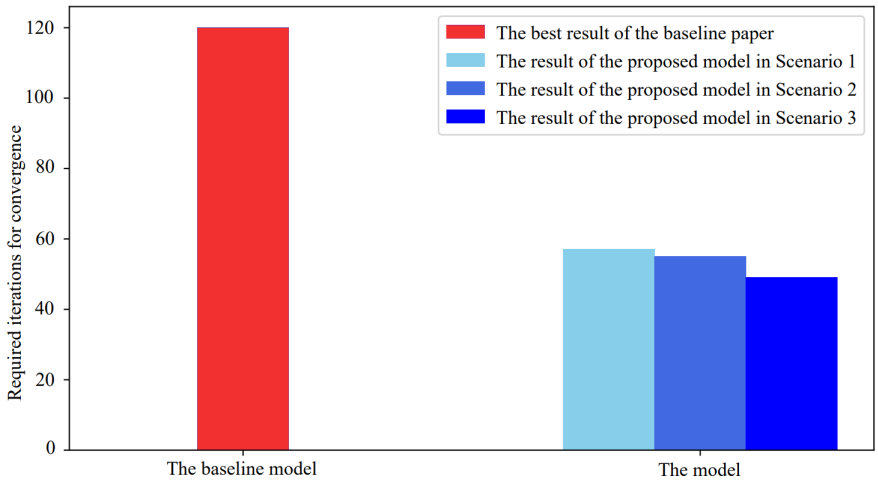


Figure 5.10 The required iterations for convergence for the proposed model and the baseline paper on a single dataset

5.7 Conclusion

The use of cobots as human assistants in disassembly projects is expanding. Cobots can undertake hazardous tasks that may jeopardize the safety of human operators. Cobots can also efficiently execute simple and repetitive tasks that human operators may not complete with high precision due to fatigue and distractions. In contrast, cobots cannot perform complex and challenging tasks that require the skill and strength of human operators. HRC disassembly significantly promises progress in the manufacturing industry by taking advantage of the flexibility of human operators and the high precision of cobots. The use of HRC in disassembly processes is a positive step towards optimizing the operations, greatly impacting recovered parts quality, resource utilization, operation time, and profit. Despite all the benefits of HRC

disassembly over traditional manual disassembly, this field still faces serious challenges. This research study aims to solve the existing challenges and fill the gaps in the literature by presenting a novel sustainable HRC disassembly planning framework based on a new multi-agent fuzzy-RL model. This model is capable of allocating tasks to operators in real-time considering the elements associated with the business aspect as well as environmental and social objectives. As a result, this leads to a sustainable disassembly process that satisfies not only business objectives but also environmental and social considerations. Given the infeasibility of completing some tasks because of the necessity for specialized tools or technology, the model incorporates the technical feasibility of tasks in the planning process. The model also appropriately represents uncertain parameters of the problem by using a fuzzy-based environment. Hence, the agent receives more accurate feedback as a reward signal from the environment. As RL agents continuously learn from reward values, this strategy enhances the learning process of the RL model thereby improving its ability to make robust decisions in diverse situations. To effectively validate the presented contributions, this research study employs a quantitative evaluation method, focusing on a case study of a hard drive disk. Furthermore, we improve the model performance by conducting sensitivity analysis and fine-tuning the impactful parameters of the problem, such as ϵ, γ, and the replay memory size. It also enables us to achieve a more profound understanding of the conditions governing the problem. The main limitation of this research is the absence of testing the proposed framework in a real industrial environment. Additionally, we assume that the energy consumed by cobots solely depends on operation. However, in real-world scenarios, additional factors such as the required movement to reach a specific location before performing a given task influence energy consumption. We will provide a sensitivity analysis based on these factors in our future research. This research also assumes that each task requires only one tool. In future research, we will consider a set of tools for each task to develop a more comprehensive planning model that more closely aligns with industrial conditions. Moreover, we aim to validate our model with a real case study in a laboratory environment. We will also present a trade-off analysis between objectives by evaluating the model's performance through varying each element coefficient in the reward functions, representing the corresponding importance in the decision-making process.

Disclosure statement

No potential conflict of interest was reported by the authors.

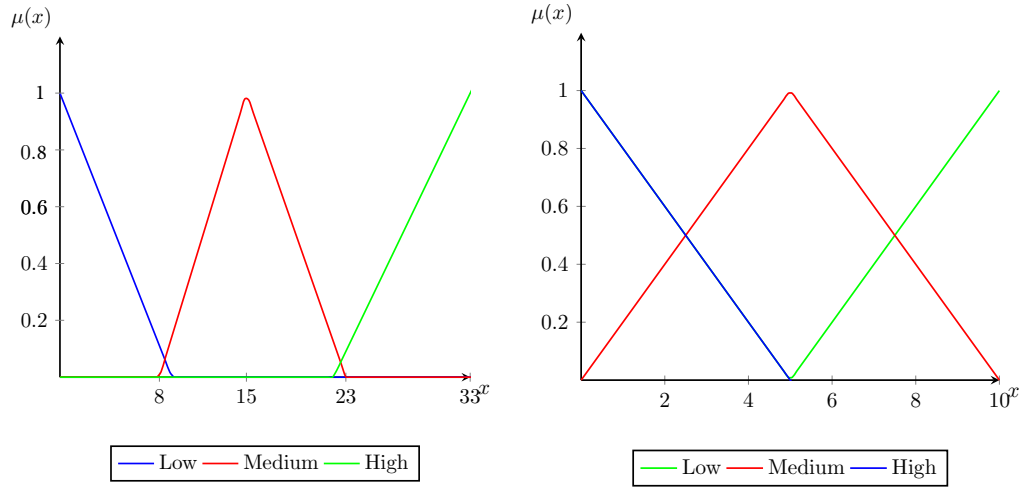
Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [grant number RGPIN-2020-05565].

Data availability statement

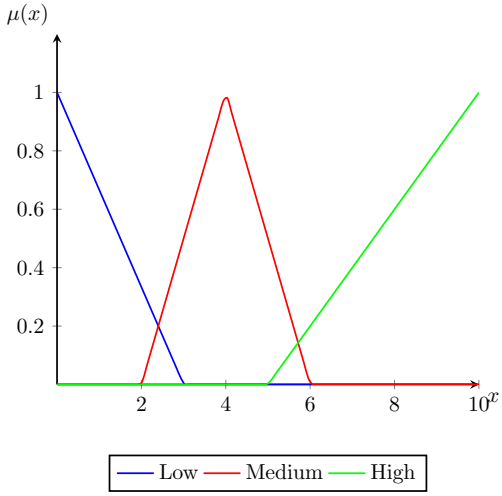
The data that support the findings of this research study are available from the corresponding author, with reasonable request.

Appendix

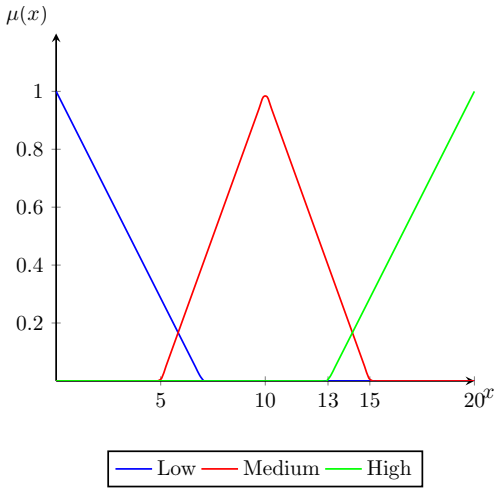


(a) Difficulty

(b) Recovered component quality



(c) Operation time



(d) Technical feasibility

Figure 5.11 The employed fuzzy numbers in the proposed model

Algorithm 5 The fuzzy-based environment modeling

For a human operator with an accepted level of experience
if (Difficulty = Low) AND (Technical Feasibility = High) **then**
    Medium ⇒ Time.
    Medium ⇒ Recovered quality.
else if (Difficulty = Low) AND (Technical Feasibility = Medium) **then**
    Medium ⇒ Time.
    Medium ⇒ Recovered quality.
else if (Technical Feasibility = Low) **then**
    High penalty.
else if (Difficulty = Medium) AND (Technical Feasibility = High) **then**
    High ⇒ Time.
    Low ⇒ Recovered quality.
else if (Difficulty = Medium) AND (Technical Feasibility = Medium) **then**
    High ⇒ Time.
    Low ⇒ Recovered quality.
else if (Difficulty = High) AND (Technical Feasibility = High) **then**
    High ⇒ Time.
    Low ⇒ Recovered quality.
else if (Difficulty = High) AND (Technical Feasibility = Medium) **then**
    High ⇒ Time.
    Low ⇒ Recovered quality.
end if
For a human operator with a medium level of experience
if (Difficulty = Low) AND (Technical Feasibility = High) **then**
    Low ⇒ Time.
    High ⇒ Recovered quality.
else if (Difficulty = Low) AND (Technical Feasibility = Medium) **then**
    Low ⇒ Time.
    High ⇒ Recovered quality.
else if (Technical Feasibility = Low) **then**
    High penalty.
else if (Difficulty = Medium) AND (Technical Feasibility = High) **then**
    Low ⇒ Time.
    High ⇒ Recovered quality.
else if (Difficulty = Medium) AND (Technical Feasibility = Medium) **then**
    Low ⇒ Time.
    High ⇒ Recovered quality.
else if (Difficulty = High) AND (Technical Feasibility = High) **then**
    Medium ⇒ Time.
    Medium ⇒ Recovered quality.
end if
For a human operator with a high level of experience
if (Technical Feasibility = Low) **then**
    High penalty.
else
    Low ⇒ Time.
    High ⇒ Recovered quality.
end if
For a cobot
if (Difficulty = Low) AND (Technical Feasibility = High) **then**
    Low ⇒ Time.
    High ⇒ Recovered quality.
else if (Difficulty = Low) AND (Technical Feasibility = Medium) **then**
    Low ⇒ Time.
    High ⇒ Recovered quality.
else if (Difficulty = Low) AND (Technical Feasibility = High) **then**
    High penalty.
else if (Difficulty = Medium) **then**
    High penalty.
else if (Difficulty = High) **then**
    High penalty.
end if

CHAPTER 6 ARTICLE 3: REAL-TIME VIDEO PROCESSING IN FUZZY POSTURE-BASED ERGONOMIC ANALYSIS IN A DISASSEMBLY CELL

Authors: Ashkan Amirnia, Elham Ghorbani, and Samira Keivanpour

Submitted on January 27, 2024, and published on August 28, 2024, in the *6th International Conference on Intelligent and Fuzzy Systems (INFUS 2024)*, pages 247-256, DOI: 10.1007/978-3-031-67192-0_31.

6.1 Abstract

Traditional ergonomic evaluations often overlook the dynamic and uncertain nature of human movements, leading to potential musculoskeletal disorders (MSDs) and impacting worker health, efficiency, and company costs. Disassembly cells, crucial for sustainability and circular economy efforts, pose unique challenges and opportunities for ergonomic optimization. This study introduces an innovative approach for ergonomic risk assessment in the manufacturing industry, particularly within disassembly cells, by integrating real-time video processing and fuzzy logic. Our research fills a significant gap in ergonomic assessment by utilizing a multi-camera computer vision technique to capture and analyze worker motions in real-time, allowing for dynamic ergonomic risks assessment in a disassembly cell. The fuzzy logic inference enhances the system's ability to handle the variability and subjectivity of human posture, offering a more nuanced and accurate risk assessment than binary logic systems. Experimental validation in a laboratory setting confirms the feasibility of our approach, demonstrating its potential to improve worker safety and productivity by providing a more responsive and adaptable tool for ergonomic assessment in industrial environments. This work marks a significant advancement in the field, suggesting a path forward for the development of ergonomic interventions that are both more effective and applicable in diverse manufacturing settings.

Keywords: Ergonomic risk; disassembly cell; fuzzy logic; computer vision

6.2 Introduction

The emergence of complex manufacturing processes necessitates the development of advanced ergonomic analysis techniques to ensure worker safety and optimize productivity. Ergonomic risk is a major concern in the manufacturing industry as it directly impacts the health and well-being of workers, resulting in potential MSDs which are a leading cause of work-related

injury and long-term disability. High ergonomic risk can result in decreased worker efficiency, increased absenteeism, and higher healthcare and compensation costs. Traditional ergonomic assessments often rely on static postures, failing to account for the dynamic nature of human movements. Furthermore, existing models typically do not accommodate the inherent uncertainty and variability of human posture and the subjective nature of discomfort and strain assessments.

A disassembly cell is a system that separates the components of a product that has reached the end of life, so that they can be reused, repaired, remanufactured, or recycled. It can involve human operators, robots, or a combination of both, working together to perform the disassembly tasks. The disassembly operation improves sustainability and circularity, because it reduces the waste and environmental impact of discarding products and preserves the value and quality of materials. By using a disassembly cell, a business can also fine-tune its operations, by optimizing the disassembly sequence, reducing the costs and time of disassembly, and improving the safety and ergonomics of the workers.

This study addresses critical gaps in current ergonomic assessment methodologies by integrating real-time video processing with a fuzzy logic inference engine to mitigate ergonomic risks in a disassembly cell. By using different computer vision techniques, we initially extract foreground as well as the magnitude and angles of motions in each frame. We then remove the corresponding noise by applying basic image processing methods, including morphological operations and thresholding. After detecting the operators' locations, we extract the geometric angles related to the orientation of their bodies. Furthermore, this approach defines a region of interest for processing based on the magnitude of motion. In this way, the algorithm avoids processing pixels corresponding to points with less movement, such as walls and windows, which are not important in this context. In addition, we down sample video frames by processing only one out of every two frames. These techniques effectively increase the speed of the proposed approach, meeting the requirements of real-time. The extracted features are then utilised to compute ergonomic risk using fuzzy methods. By employing a fuzzy logic inference engine, the system evaluates these risks with a delicate approach that mirrors human reasoning. The fuzzy logic enables the proposed system to handle imprecise input data, such as the variability of human posture, and provide a more accurate assessment of ergonomic risks than binary logic systems.

To validate the proposed framework, experimental analysis has been conducted in a laboratory environment. The integration of real-time video processing allows for continuous monitoring and analysis of worker postures, leading to a proactive approach in identifying and mitigating potential MSDs. The fuzzy logic component addresses the research gap of

quantifying subjective ergonomic factors by providing a flexible and adaptable model that can be tailored to various industrial settings.

The anticipated outcome of this research is a robust ergonomic analysis tool that enhances the safety and efficiency of disassembly cell design. By addressing the limitations of current ergonomic assessment tools and incorporating real-time video analysis with fuzzy logic, this study aims to pave the way for more responsive and adaptive ergonomic interventions in the manufacturing industry. The experimental validation demonstrates the practicality of the proposed system, marking a significant step forward in real-time ergonomic risk assessment.

The structure of this manuscript is as follows: Section 6.3 explains the problem context. Section 6.4 delineates the proposed framework. In Section 6.5, the practical implications of the proposed framework are discussed. Finally, Section 6.6 presents concluding remarks.

6.3 Problem context

6.3.1 Product disassembly

Product disassembly involves separating components from an end-of-life (EoL) product with the purpose of reuse in remanufacturing processes. Product disassembly is an important phase in remanufacturing and circular economy by reducing the need for new raw materials, yielding considerable financial and ecological achievements. Disassembly planning refers to generating an optimum sequence of disassembly tasks to minimize costs in addition to maximizing profits. A considerable body of literature has investigated various rule-based and learning-based models to plan disassembly processes. Zhou et al. comprehensively review disassembly planning methods, various approaches for modeling disassembly processes, existing challenges, and future perspectives [2]. In [133], the authors developed a classical reinforcement learning (RL) model for disassembly planning. The associated reward function includes terms related to operation time, changes in the process direction and tool usage, the volumes of parts, and worn components. The model also integrates precedence relations among parts in the decision-making process. Chen et al. [46] developed a Q-tabular model for the disassembly planning of EoL smartphones. Considering the constraint relationship among components, the model aims to minimize the total time, which is deterministically defined as the sum of times required for tool changes, basic disassembly operations, positioning, and cleaning. [143] introduces a hybrid planning model based on fuzzy logic and augmented reality for aircraft disassembly. A recent review of the literature on this topic reveals that employing collaborative robots (cobots) as human assistants in disassembly processes is becoming more and more popular. Human-robot collaboration (HRC) effectively

combines the cobots' preciseness with the flexibility and strength of humans. This strategy has great potential to improve the profitability and efficiency of disassembly processes. There is a vast amount of literature on developing optimization models for HRC disassembly planning. In [140], the authors have proposed a real-time HRC disassembly planning model based on a dual-agent RL algorithm. The proposed model allocates tasks to a human operator and a cobot in real-time, enabling the model to overcome uncertainties that may change a process's expected flow. The model also integrates different features of parts and human skill levels in the decision-making process. [86] models a product architecture and constrain relations between parts through an AND/OR graph in an HRC product disassembly. The authors categorized disassembly tasks into eight classes according to corresponding complexity and difficulty. This research optimizes the process by using a genetic algorithm (GA) with a fitness function including operation time, human operators change frequency, and a term about the selection of non-profitable parts. [128] introduces an HRC disassembly planning approach that involves partial destructive operations. This approach considers the product's failure factors and calculates disassembly modes, as well as the values of recovered parts. The authors optimized the process with an extended GA approach based on operation time, physical constraints among components, product failure factors, disassembly difficulty, and cost.

6.3.2 Ergonomic consideration in disassembly cells

A disassembly process may involve tasks that pose high ergonomic risks. Examples of such tasks include lifting heavy loads, performing destructive operations requiring a high degree of force, maintaining an inappropriate position for a long period, and frequently doing simple and repetitive activities such as removing screws. Engaging in these tasks may result in serious body skeleton issues for humans in the long run. It is necessary to assess the ergonomic risk associated with each task and use cobots, assistant robots, wearable devices, or any advanced tools to assist humans in performing high-risk tasks. Furthermore, it is crucial to integrate ergonomic risk within the disassembly planning process. The sequence of disassembly tasks should be planned to minimize ergonomic risks for an operator. Many attempts have been made aimed at analyzing ergonomic risks in assembly/disassembly processes. [72] investigates ergonomic risk in manual and HRC disassembly processes. It experimentally shows that the synergy between a human and a cobot enhances ergonomics relative to manual disassembly. Siew et al. [139] introduced an approach to improve ergonomics in disassembly and maintenance processes. The proposed approach optimizes the disassembly process by considering real time ergonomic feedback based on data, which are collected from a sensor network. Although significant strides have been made to improve disassembly processes, few

researchers have addressed the problem of ergonomics in disassembly cells. Consequently, the investigation of ergonomic factors in this context is a new research area.

6.3.3 Research gaps in ergonomic assessment tools

Observational techniques such as rapid upper limb assessment (RULA) [144] and rapid entire body assessment (REBA) [145] have notable limitations when it comes to assessing ergonomic risks. Firstly, the evaluation of risk levels typically requires experienced assessors, making these methods potentially less cost-effective. Secondly, due to the inherent subjectivity of evaluators, the final score can suffer from inconsistency [146]. Thirdly, current ergonomic assessment tools (EATs) typically consider multiple body parts, such as the upper limbs in RULA or the entire body in REBA, and integrate the risk scores of each to anticipate overall risk levels for each task. However, in some manufacturing applications such as disassembly cells planning or job rotation approaches in assembly lines, understanding the total risk associated with each specific body area can be more conducive to implementing preventive ergonomic interventions. Given that observational EATs are prone to human error, resulting in outcomes with low consistency and repeatability, this paper introduces a comprehensive fuzzy assessment tool designed to address the imprecision inherent in traditional EATs. Moreover, to the best of our knowledge, conventional EATs evaluate postural risk by roughly considering load and time factors. For example, in REBA, the force parameter is evaluated based on 5 or 10 kg thresholds, while in RULA, task duration is only accounted for if it exceeds one minute or repeats more than four times per minutes, with a single score added to the risk in these cases. However, incorporating task time as a percentage of cycle time (CT) provides a clearer understanding of the frequency and repetition of each task. Furthermore, the load placed on various joints, or the force exerted on different muscles can be more accurately assessed using metrics such as maximum voluntary contraction (MVC) [147] or maximum acceptable effort (MAE) [148]. Additionally, the influence of time parameters in exacerbating joint stress can be better evaluated within the context of CT, which more effectively captures task repetition and resulting fatigue [149]. Therefore, this study aims to address current research gaps by proposing a fuzzy EAT that evaluates cumulative ergonomic risk for each body part, taking into account task duration as a percentage of CT and incorporating load parameters in the form of % MVC for each body part.

6.4 Proposed fuzzy approach

Figure 6.1 illustrates the outline of this research study. First, a multi-camera vision approach detects the joints of a human operator working in a disassembly cell, then angles between the

joints are computed. Finally, we analyze ergonomic risks by using a fuzzy-based approach based on the angles.

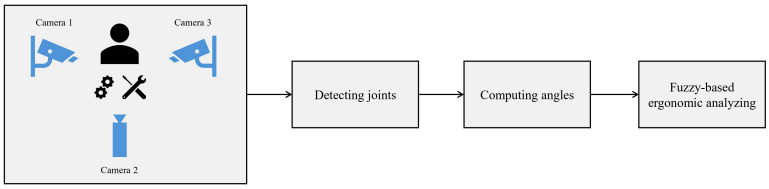


Figure 6.1 The proposed vision-based system for analyzing ergonomic risks

6.4.1 Fuzzy ergonomic assessment tool

In this subsection, we present a comprehensive method for evaluating ergonomic risks using fuzzy logic. Our novel approach integrates posture risk and fatigue levels of specific body parts, assessing cumulative ergonomic risk based on task duration relative to CT. Initially, we focus on two body parts—upper arm and lower arm— to prototype a model for future studies. The following subsections detail our proposed model in three steps to elucidate its underlying logic.

Posture risk analysis To assess posture risk for each body part, we employ the REBA and RULA methods, interpreting risk levels akin to a traffic light scheme (green for low risk, yellow for medium risk, and red for high risk). Tables 6.1 and 6.2 present primary fuzzy rules for posture risk assessment, considering various joint angles. These tables utilise side videos to estimate x angles and apply up-hand side videos to capture y angles. Negative angles denote positions behind and inclined to the center of the body, on x and y sides, respectively.

Fatigue consideration By incorporating MVC and MAE from previous studies [147, 148], we establish nine fuzzy rules (Table 6.3) linking the percentage of MVC (or MAE) to posture risk aiding in fatigue assessment.

Time-based cumulative evaluation For assessing cumulative ergonomic risk, a fuzzy

Table 6.1 Posture risk of upper arm (shoulder)

Angle from the core	$y \leq 0°$	$0° < y \leq 20°$	$20° < y \leq 45°$	$45° < y \leq 90°$	$90° < y$
$-20° \leq x < 0°$ or $0° < x \leq 20°$	1 & 0	1 & 1	1 & 2	1 & 3	1 & 4
$x < -20°$ or $20° < x \leq 45°$	2 & 0	2 & 1	2 & 2	2 & 3	2 & 4
$45° < x \leq 90°$	3 & 0	3 & 1	3 & 2	3 & 3	3 & 4
$90° < x$	4 & 0	4 & 1	4 & 2	4 & 3	4 & 4

Table 6.2 Posture risk of lower arm (elbow)

Angle from the core	$y = 0°$	$0° < y$ or $y < 0°$
$60° \leq x \leq 100°$	1 & 0	1 & 1
$x < 60°$ or $100° < x$	2 & 0	2 & 1

Table 6.3 Integrating fatigue parameter to posture risk.

Load	MVC% $\leq$ 15%	15% $<$ MVC% $\leq$ 40%	40% $<$MVC%
Posture	Low	Medium	High
Low	Low	Low	Medium
Medium	Low	Medium	High
High	Medium	High	High

expert system generates rules based on task durations relative to CT. These fuzzy rules, informed by ergonomic expert insights, guide risk categorization:

1. If no high-risk tasks are present and cumulative medium-risk task time exceeds 50% of CT, then the operator faces medium risk (orange).
2. If no high-risk tasks are present and cumulative low-risk task time exceeds 50% of CT, then the operator faces low risk (green).
3. If cumulative high-risk task time exceeds 20% of CT, then the operator faces high risk (red).
4. If cumulative high-risk task time falls between 0 and 20% of CT, and the cumulative low-risk task time exceeds 50% of CT, then the operator faces minor risk (yellow).
5. If cumulative high-risk task time falls between 0 and 20% of CT, but cumulative low-risk task time is less than 50% of CT, then the operator faces medium risk (orange).

6.4.2 Real-time video processing method

This section discusses the computer vision method employed in this research. This method aims to accurately measure the angles between the joints of a human operator, which is working in a disassembly cell. Since it is impossible to extract all joints information from one camera, we have developed a multi-camera system utilizing three cameras. As illustrated in Figure 6.1, these cameras are installed on the right, front-facing, and left sides of the human operator, respectively. This approach enables us to effectively extract all the required information.

This research study initially extracts the joints in each frame of the videos captured by the three cameras. It uses the Media pipe library, which is an open-source framework provided by Google, for joint extraction. Since this research focuses on the ergonomic risk associated with the upper body, we specifically extract the wrist, elbow, shoulder, and hip joints from videos. In this context, we need to compute the angle between elbow, shoulder, and hip (referred to as angle 1) and the angle between wrist, elbow, and shoulder (referred to as angle 2). We use Formula 6.1 to calculate the angles. Figure 6.2 depicts the vectors v_1 and v_2, along with the angle θ. To calculate angle 1, we determine the vectors v_1 = (shoulder, elbow) and v_2 = (shoulder, hip). Similarly, we define vectors v_1 = (shoulder, elbow) and v_2 = (elbow, wrist) to calculate angle 2. In the following, these angles are used as the input of the fuzzy model.

$$\text{Angle between vectors } (v_1, v_2) = \arccos\left(\frac{v_1 \cdot v_2}{\|v_1\| \cdot \|v_2\|}\right) \tag{6.1}$$

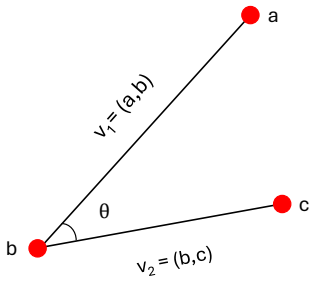


Figure 6.2 Two vectors and the corresponding angle between them

6.5 Application perspective

Utilizing the proposed method in disassembly cells offers notable advantages in real-world scenarios due to its capacity to detect cumulative ergonomic risk in each body part and help decision-making process for applying any ergonomic interventions. The proposed framework can include more body parts, thereby facilitating taking proper action in improving the ergonomic level of the workplace. This approach can be applicable across diverse industries and assists them in finding robust solutions. To explain the effectiveness of this methodology, a small numerical example is presented. This example is for the first five tasks of disassembly of a standing fan (as shown in Figure 6.3). The corresponding angles of the two body parts (shoulders and elbows) are displayed in Table 6.4 and based on Table 6.1 and 6.2 the posture risk levels are shown by color codes. Then, by considering the load of each task and applying

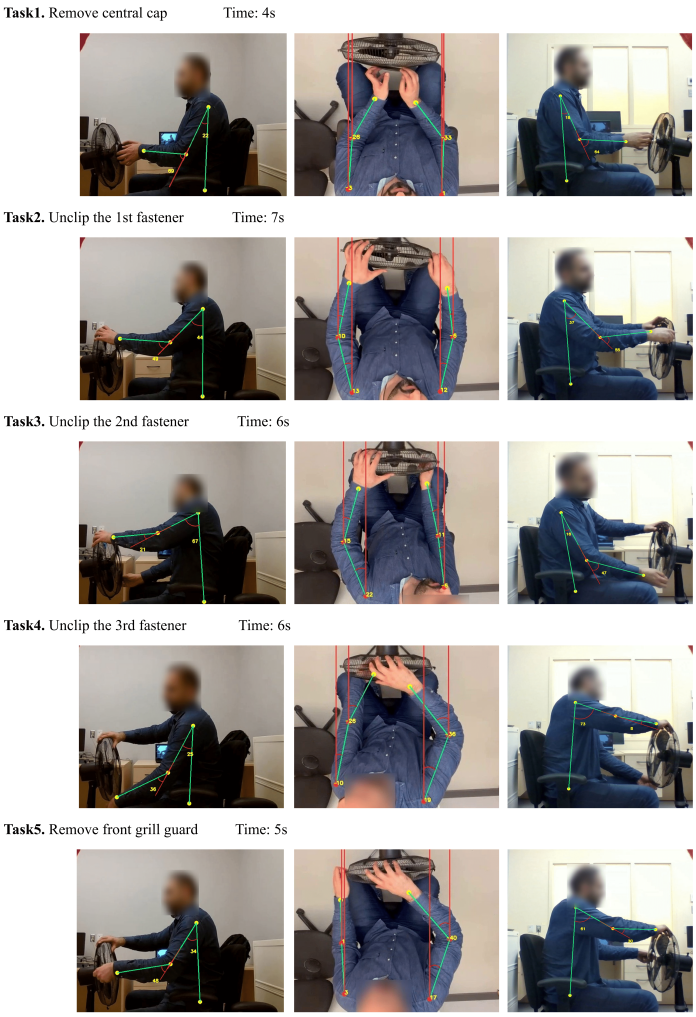


Figure 6.3 Detected joints and computed angles in the utilised case study. The collected image data was used for initial feasibility testing and was not subject to detailed analysis in this study

the fuzzy rules depicted in Table 6.3, the risk level of each task in each body part is detected (see Table 6.5). Finally, based on the duration of each task and comparing it by CT (equal to 35 s in this example), fuzzy rules are employed to determine the cumulative risk of each joint as presented in the last row of Table 6.5.

6.6 Conclusion

Product disassembly involves tasks such as performing simple and repetitive activities for extended periods, moving heavy objects, and carrying out dismantling operations. These tasks may present different ergonomic risks for human operators. This paper presents a

Table 6.4 Posture risks based on angles detected for disassembly tasks.

	Angles	Left Shoulder		Left Elbow		Right Shoulder		Right Elbow	
Task		x	y	x	y	x	y	x	y
1		22	-3	69	-26	18	1	64	-33
2		14	13	49	-10	37	12	55	-6
3		67	22	21	-15	16	-6	47	-11
4		25	-10	36	-26	73	19	8	-36
5		34	3	48	1	61	17	30	-40

Table 6.5 Final and cumulative risks of each body part.

Task	Loads	Left Shoulder	Left Elbow	Right Shoulder	Right Elbow	Time/CT
1		15%	19%	17%	10%	11%
2		18%	20%	10%	12%	20%
3		10%	25%	25%	15%	17%
4		42%	9%	42%	20%	17%
5		12%	10%	14%	16%	14%
Total Risk		Rule 2	Rule 1	Rule 4	Rule 3	

novel method for ergonomic risk assessment in a disassembly cell, utilizing fuzzy logic and computer vision. The method starts by extracting the human upper body joints by processing the videos captured from three cameras installed in the cell. Subsequently, the angles between the joints are calculated. A fuzzy-based approach is then utilised to analyze ergonomic risk according to the acquired angles. This research also presents a real case study conducted in a disassembly cell to validate the proposed method. The obtained results demonstrate the method's effectiveness in analyzing ergonomic risks associated with disassembly tasks. In our future research, we will embed this fuzzy ergonomic approach into all human joints and implement the proposed method in real world settings to examine the validity of it by comparing the results with biomechanical and exact tools. In addition, future research will provide a more comprehensive validation by involving multiple participants with different characteristics, such as gender, height, and weight, in experimental settings.

CHAPTER 7 ARTICLE 4: A FUZZY-GUIDED REINFORCEMENT LEARNING METHOD FOR SUSTAINABLE COBOTIC DISASSEMBLY PLANNING UNDER UNCERTAINTY

Authors: Ashkan Amirnia and Samira Keivanpour

Submitted on January 16, 2025, in the *International Journal of Production Research (IJPR)*, Submission number: 255194371.

7.1 Abstract

End-of-life (EoL) product disassembly is an important step in remanufacturing, resulting in significant economic and environmental benefits. Traditionally, disassembly operations are carried out manually, which is costly, time-consuming, and requires skilled human operators. Furthermore, output quality highly relies on operators' skills. Transitioning from manual disassembly to automated disassembly by employing collaborative robots (cobots) as human assistants has significant potential to improve process efficiency and quality. Cobots can effectively perform simple and repetitive tasks as well as dangerous tasks that pose high risks to human health. Moreover, it is crucial to optimize disassembly processes by generating a task sequence for each operator to maximize profit and minimize cost. This study presents a novel multi-agent approach based on reinforcement learning (RL) to optimize human-robot collaboration (HRC) disassembly processes. In addition, it involves experts' knowledge within the RL model by developing a fuzzy-based exploration method. An electronic board is used as a case study to validate the proposed approach. The results indicate that the model using fuzzy exploration achieves more stable and sustained reward values. This research also conducted multiple sensitivity analyses, proving the model's capability to make real-time decisions in uncertain scenarios based on the different importance of sustainability objectives.

Keywords: Disassembly planning; reinforcement learning; fuzzy logic; human-robot collaboration; sustainable driven planning

7.2 Introduction

Nowadays, population growth, rising consumerism, and the shift towards modernity have significantly increased the production of goods. This has resulted in the generation of vast amounts of EoL products each year, posing serious environmental concerns. Landfilling, incineration, recycling, and remanufacturing are among the most common solutions to address

this challenge. However, landfilling or incineration pollutes the air, soil, and water. While recycling is a more eco-friendly solution, it may consume considerable amounts of energy. On the other hand, remanufacturing focuses on rebuilding EoL products using various methods instead of discarding, burning, or recycling them [2]. A key step in remanufacturing is product disassembly, which refers to the process of separating a product into its components. This procedure offers significant environmental and economic benefits, such as waste reduction, lower costs of waste management, preserving natural resources and raw materials, and energy footprint mitigation.

Conventionally, disassembly operations are manually executed by labours, a time-intensive and expensive approach. In recent years, the use of cobots has become more and more popular in industrial processes, such as product disassembly. Cobots can accurately execute repetitive and easy tasks that human operators may not carefully complete due to fatigue or distraction. Cobots can also perform dangerous tasks, such as tasks requiring exposure to toxic substances or tasks with high ergonomic risks. Conversely, cobots lack the flexibility and strength to perform complex and difficult tasks. On the other hand, human operators can effectively execute such complex and heavy tasks. Hence, the collaboration of humans and cobots leverages both the flexibility of humans and the accuracy of cobots, offering the potential to improve the productivity and efficiency of disassembly processes.

Although involving cobots in such processes may enhance efficiency and quality, it is crucial to consider the amount of energy consumed by cobots. The increased use of cobots causes more energy consumption, which not only increases costs but also generates excessive energy-related pollutants. Furthermore, performing complex and heavy tasks with cobots could potentially damage the product or cobots physical structures. It is also critical to consider human factors such as ergonomic risks and human safety while allocating tasks to operators. Thus, optimized task assignment is significantly important as inefficient allocating tasks may greatly reduce process productivity. Disassembly planning is an essential part of the management of EoL products, addressing the sequential assignment of disassembly tasks to operators to minimize costs and maximize profitability. In addition, an important step in planning a process is modeling a product structure, which refers to representing precedence relationships between components of the product [2].

A disassembly process involves uncertain factors that potentially change the process from its expected routine. These factors include human errors, variable product conditions, machine failure, tool tear and wear, and any other external disruptions. These factors may result in the prolongation or even failure of operations. Hence, it is not practical to use static planning models that theoretically generate a pre-planned task sequence. A disassembly

planning model should dynamically allocate tasks to operators based on online conditions to cope with uncertainties.

There is a considerable amount of literature on developing optimization methods, such as the Genetic algorithm (GA), discrete Bees algorithm, and artificial bee colony (ABC), to plan processes based on different criteria. Moreover, different graph-based and matrix-based approaches are used to model precedence relations within parts of products. [86] presents disassembly precedences with an AND/OR graph-based approach. Then, the GA optimizes task allocation between a human operator and a cobot, considering several parameters such as operation and tool change times, task difficulty, and operator change. Similarly, Lu et al. [150] modeled disassembly precedence using an AND/OR approach. They then planned the process based on profit and energy with a hybrid metaheuristic algorithm. [59] proposes a multi-objective optimization approach based on ABC algorithm to plan a fully robotic disassembly process concerning makespan and energy consumption. The proposed approach considers three phases of energy consumption: disassembly operation, standby, and tool change. [39] optimizes a robotic disassembly planning problem with a proposed multi-objective robust approach. It aims to cope with uncertainties regarding products and operator conditions during a process.

Xu et al. [151] proposed a matrix-based approach to extract the precedence relationships of components in a product. Subsequently, a modified discrete Bees algorithm based on Pareto optimizes the HRC disassembly process concerning disassembly time, cost, and difficulty. In this case, the difficulty scores for humans and robots are evaluated separately. For humans, these scores depend on workload and the level of hazards involved in tasks, whereas difficulty scores are determined by movement complexity, operation space, and perception capability for robots. [128] introduces an improved GA to optimize an HRC partial destruction disassembly process. The proposed model considers several characteristics in the planning process. These characteristics are disassembly difficulty, total disassembly time, cost, and parts failure information. In a recent study, [73] plans the HRC process by a hybrid ant lion optimizer according to multiple objectives, including the ergonomic consideration of humans, parts' recycling revenue, disassembly complexity, and operation time. The authors used a video processing approach to assess the ergonomic aspect. This approach extracts human joints from video frames to calculate the angles between the joints. Then, the rapid entire body assessment (REBA) approach measures the ergonomic factor based on the angles. In addition, the study models the process with an AND/OR graph approach.

In order to provide a sustainable and reliable disassembly process, this research presents a learning-driven approach with real-time decision-making capability. The major contributions

of this study are as follows.

- This paper presents a novel disassembly planning approach based on RL and fuzzy logic that dynamically plans the process based on real-time conditions instead of relying solely on preplanned task sequences. It dynamically makes decisions after each operation, whether successful or not, allowing it to adapt to uncertainties.
- This article enhances a conventional RL model by integrating a novel fuzzy exploration method, configured based on experts' knowledge, into the RL model structure.
- Evaluate the proposed approach on a dataset and provide an in-depth comparative analysis between different exploration-exploitation scenarios.
- Present new sensitivity analyses to assess the model's performance in uncertain scenarios, such as variable execution times and a failure probability for each operation. This article also provides a trade-off analysis between sustainable objectives in decision-making.
- This research develops a graphical user interface (GUI) that enables users to easily customize the planning of disassembly processes according to required sustainability objectives.

The rest of this article is organized as follows. Section 7.3 is dedicated to review the literature. Section 7.4 introduces the problem of optimizing HRC disassembly task sequences and the associated economic, environmental, and social objectives. Section 7.5 presents our proposed fuzzy-RL model to optimize HRC disassembly processes. Section 7.6 provides the experimental results and discussion. Finally, Section 7.7 concludes the article.

7.3 Related works

A growing body of literature has explored various rule-based and learning-based data-driven disassembly planning methods. Section 7.3.1 reviews the existing methods based on classical optimization algorithms, while Section 7.3.2 discusses learning-based models. Following this, Section 7.3.3 presents a qualitative comparison between our proposed approach and current methods in the literature, highlighting its novelty. The comparison focuses on the disassembly features included in the planning process.

7.3.1 Classical planning models

Traditionally, classical optimization methods have been employed to plan disassembly processes, and they have been widely investigated by researchers. For instance, [152] proposes an improved discrete Bees algorithm with variable neighborhood search for balancing robotic disassembly lines. A matrix-based approach is also used to generate feasible disassembly sequences and directions. The authors evaluated the proposed approach using a gearbox and a camera as case studies. Guo et al. [77] introduced a partial destructive disassembly planning approach to address the line balancing problem. This approach considers the feasibility of destructive operations and the resulting noise pollution during the planning process. It presents a mathematical model with different objectives, including minimizing the number of stations, noise pollution, and related costs. Subsequently, an enhanced gray wolf algorithm is developed to optimize the task sequence. Zhong et al. [16] introduced a disassembly planning approach that incorporates Dijkstra's algorithm and PSO, considering both functional components and fasteners. [33] proposes a disassembly planning approach that uses a graph-driven method to represent product architecture. Then, it optimizes the task sequence using a hybrid model based on GA and Tabu search.

In a recent study, [35] develops an approach based on an improved ABC algorithm for balancing partial multi-robotic disassembly lines. The proposed approach considers profitability, cycle time, energy consumption, and the extra time and cost resulting from workstations' reconfiguration. [31] optimizes task sequences using the ant colony optimization (ACO) algorithm with respect to disassembly directions, component numbers, and tools. [36] proposes a multi-objective approach to balance position-constrained HRC disassembly lines. They presented a mixed integer programming model and a multi-fidelity optimization algorithm to address both small-scale and large-scale problems, respectively. Zhan et al. [153] developed a planning model based on the Northern Goshawk algorithm for vehicle battery disassembly, with the main objectives of minimizing energy consumption and hazards during the planning process. [37] presents a hybrid approach based on general ontology and rule-based reasoning for HRC disassembly planning. They validated the proposed approach by using a gearbox as a case study. In a more recent research study, Lou et al. [38] have developed a human-cyber-physical system framework for HRC disassembly planning, in which the role of humans is considered in two human-in-the-loop (HitL) and human-on-the-loop (HotL) paradigms. The authors also computed task complexity and operator ergonomics using a cloud-based approach. Finally, an enhanced hybrid grey wolf optimization algorithm plans the sequential task allocation process. They conducted an experimental analysis by applying the proposed framework to a control box as a real case study.

7.3.2 Learning-based planning models

In recent years, with the growing rate of machine learning (ML) applications in various domains, many attempts have been made in order to optimize disassembly processes by adopting ML models. In particular, RL methods are becoming increasingly popular among researchers in this field. [48] optimizes task sequences in the disassembly process using a graph-based RL approach. It categorizes components into different levels according to their precedence relations and position on a graph, which represents a product structure. The proposed approach begins by selecting components at the highest level and proceeds to the next level after disassembling all components at each level. The RL model selects the next component for disassembly to maximize a reward value that is a linear combination of several objectives, including operation time, profits, tool and direction change, in addition to a penalty term. This term incentivizes the model to select the components in line with the current level. In a more recent study, [140] models a product structure through a graph-based approach. They also developed a planning approach based on a multi-agent RL model to optimize the HRC disassembly process with respect to the characteristics of the products and operators.

Mao et al. [45] implemented their proposed planning approach into a virtual reality (VR) framework for maintenance training applications. The approach represents a product structure through a Petri Net. It also comprises a hybrid GA-RL model that optimizes the task sequence concerning difficulty, time, and cost. By developing a hybrid model based on PSO and Q-learning, [127] optimizes the HRC disassembly process to achieve minimum operation time. In each evolutionary state of PSO, the Q-learning approach selects the most effective optimizer for improving performance. In addition, the authors used multiple perturbation and local search techniques to enhance the planning process. Allagui et al. [133] extracted precedence relationship information from a CAD model of a product to compute all feasible disassembly sequences along the x, y, and z axes. They then represented mutual connections between components with a square matrix, whose dimensions equal the number of the product's components. This matrix has a value of -1 for each index between two components that are not connected. It has a value of +100 for the indexes corresponding to connected components. Additionally, the indexes on the diagonal axis of the matrix have 0 values. Subsequently, a Q-tabular model optimizes the process with a reward function including the feasibility of movement toward the axes and a fitness function presenting the problem's objectives and characteristics. They are operation quality and time, components' volumes, and changes in tool and process direction.

Recently, Tabar et al. [52] have proposed an approach for disassembly planning that models

a process with a Petri Net. It optimizes the process by a Q-learning method. The authors considered three criteria of time, quality, and process capability as the main objectives in the reward function. In addition, they evaluated the model with an engine starter motor. [90] introduces a theoretically cyclic framework for HRC disassembly planning, which includes five steps. It starts by fusing data captured from different sensors, such as Microsoft Kinect, Ultrasonic, and Leap Motion. Subsequently, it extracts multiple types of high-level information from the data, including the human body skeleton, tool recognition, hand gestures, and operators' locations. By using this information, an RL model allocates tasks to cobotic and human operators. After completing a task and assessing the results, the acquired knowledge by the model is transferred through a cloud-based approach to other models. This process is recursively repeated until all tasks are completed.

7.3.3 Comparison of the proposed approach with existing methods

Synthesis on the literature review highlights that a key problem with much of the literature is that most existing models plan the process statically, unable to effectively perform in uncertain scenarios, where an operation takes longer time or even fails. In addition, sustainable aspects are not well addressed in this context. Few studies involved sustainable objectives in planning the disassembly process. There are potential gaps in the literature for a comprehensive sustainable planning model and a trade-off analysis between sustainability objectives during decision-making.

Table 7.1 presents a qualitative comparison between our proposed approach and the significant studies from the literature. It compares key features of the disassembly process addressed in planning models. Although each previous method incorporated some of these features, the approach proposed in this article considers a broader range of features and provides a more comprehensive and sustainable planning process.

Table 7.1 A qualitative comparison of the proposed approach with the recent literature

	HRC	Real-time planning	Difficulty	Time	Circularity	Energy	Expert knowledge	Tool change	Ergonomic risk	Safety
[86]	√	×	√	√	×	×	×	√	×	×
[150]	×	×	×	√	√	√	×	×	×	×
[59]	×	×	×	√	×	√	×	√	×	×
[39]	×	×	×	√	×	√	×	√	×	×
[151]	√	×	√	√	×	×	×	√	×	×
[128]	√	×	√	√	√	×	×	√	×	×
[73]	√	×	√	√	√	×	×	×	√	×
[38]	√	×	√	√	×	×	×	√	√	√
[48]	×	×	×	√	×	×	×	√	×	×
[140]	√	√	√	√	√	×	×	×	×	×
[45]	×	×	√	√	×	×	×	√	×	×
[127]	√	×	√	√	×	×	×	√	×	×
[133]	×	×	×	√	×	×	×	√	×	×
[52]	×	×	×	√	√	×	×	×	×	×
[90]	√	×	×	√	×	√	×	√	×	√
[60]	×	×	×	×	×	√	×	√	×	×
[88]	√	×	×	√	×	×	×	×	×	×
[1]	√	×	×	√	×	×	×	√	×	√
Our model	√	√	√	√	√	√	√	√	√	√

7.4 Problem statement

HRC disassembly planning is a multi-objective optimization problem where a model allocates tasks to human(s) and cobot(s) to minimize costs and maximize profits. These costs and profits result from the sum of various factors, defined according to the sustainability criteria (Section 1.2) and explained in detail below. Equation 7.1 shows the problem's main objective: minimizing total costs and maximizing overall profits. In this case, f_p presents the cost of element p, and g_q denotes the profit of element q. In addition, P and Q indicate the number of cost and profit elements, respectively.

$$\text{Minimize} \left(\sum_{p=1}^{P} f_p - \sum_{q=1}^{Q} g_q \right) \tag{7.1}$$

Operation time is a critical economic factor in disassembly processes. Longer operation times reduce the productivity and profitability of manufacturing companies. Hence, the planning model should consider minimizing the operation time as an objective. Equation 7.2 indicates the objective of minimizing operation time, in which T_o represents the overall process time and t_i represents the time required for disassembling component i. Furthermore, N refers to the total number of components.

$$f_1 = T_o = \sum_{i=1}^{N} t_i \tag{7.2}$$

During a disassembly process, a human operator may require to change a tool. As tool change is a time-consuming activity, the planning model should allocate tasks efficiently to minimize the number of tool changes. Equation 7.3 shows the objective of minimizing the change of tools, with $tool_i$ and $tool_{i+1}$ are the required tools to complete task i and task i+1, respectively. In this case, the term δ is a binary function that is zero if its inputs are equal and 1 otherwise. Furthermore, T_T represents the total number of tool changes in the process.

$$f_2 = T_T = \sum_{i=1}^{N-1} \delta \left(tool_i, tool_{i+1} \right) \tag{7.3}$$

Human operators' health in disassembly processes is an important consideration. Performing operations that expose humans to high ergonomic risk may result in serious health issues, such as musculoskeletal disorders (MSDs). Therefore, a disassembly planning model should allocate tasks ensuring that human operators are exposed to the minimum ergonomic risks. Equation 7.4 shows the minimizing ergonomic risk objective, where R_T is the total risk,

calculated as the sum of the ergonomic risk for each task i (denoted as r_i).

$$f_3 = R_T = \sum_{i=1}^{N} r_i \tag{7.4}$$

The energy consumption of cobots is a critical issue in industrial HRC processes. High energy utilization is costly and contributes to carbon emissions, which raise serious environmental concerns. Hence, it is vitally important to manage the energy consumed. An objective of a planning model is minimizing the amount of energy consumed during the process, as shown in Equation 7.5. Herein, e_i represents the energy required for completing the i-th task, and E_T denotes the total consumed energy.

$$f_4 = E_T = \sum_{i=1}^{N} e_i \tag{7.5}$$

Maximizing the quality of recovered parts is another objective of a disassembly process. Increasing the quality of these parts improves the productivity of remanufacturing processes and reduces the demand for new resources. In addition, low-quality recovered parts may be classified as waste and not used in remanufacturing processes. Equation 7.6 illustrates the objective of maximizing the quality of recovered parts, where Q_T represents the total recovery quality and q_i denotes the quality of the i-th component.

$$g_1 = Q_T = \sum_{i=1}^{N} q_i \tag{7.6}$$

The safety of human operator is one of the most significant factors in disassembly processes. Human exposure to potentially hazardous tasks may result in serious health issues. A planning model should allocate tasks to operators concerning human safety. Equation 7.7 represents the objective of maximizing safety, given that S_T and s_i identify the total safety during the process and safety associated with each task i, respectively.

$$g_2 = S_T = \sum_{i=1}^{N} s_i \tag{7.7}$$

This section addresses the objectives of the HRC disassembly planning problem. The next section discusses the developed Fuzzy-RL model for HRC disassembly planning. With respect to these sustainable objectives, this model makes decisions in real time to cope with uncertain scenarios outlined in Section 7.2. This leads to a shift toward a more sustainable disassembly process.

7.5 Methodology

In this section, we introduce our proposed approach for HRC disassembly planning. This section begins by outlining the proposed architecture and its modules, along with how a fuzzy model can enhance the performance of an RL model. Subsequently, we explain the RL models and the corresponding elements. Following that, we present the fuzzy models and related numbers and rules. The environment modeling and failure probability for each operation are then discussed. Finally, we delve into a detailed explanation of the entire proposed algorithm.

7.5.1 Main architecture

Figure 7.1 illustrates the schematic overview of the proposed fuzzy-RL approach that contains two models: one RL and other fuzzy. It also comprises a collaborative intelligence module and an environment. Similar to the classical RL concept, the RL model interacts with the environment and learns autonomously based on the feedback it receives. In addition, the fuzzy model is logically configured by experts. In each iteration, the collaborative intelligence module selects one of the two models for decision-making. In this case, the fuzzy model operates more than the RL model in the first iterations. As the learning process evolves, the collaborative intelligence module selects the RL model at an increasingly exponential rate, resulting in a declining influence of the fuzzy model. Ultimately, the RL model independently plans the process.

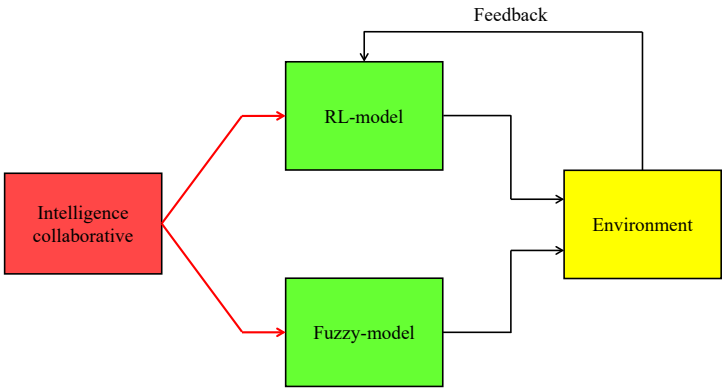


Figure 7.1 The conceptual framework for fuzzy-RL integration

In general, RL models adopt non-optimal solutions in initial interactions as they lack sufficient knowledge about the environment. They then get closer and closer to choosing the optimal solutions with the continuation of the learning process. In our proposed approach, to enhance the RL model performance, the fuzzy model directs it in the initial interactions.

Figure 7.2 provides a more detailed view of the fuzzy-RL method. The RL model interacts with the environment by selecting an action a_t at each state s_t. Subsequently, it receives a reward r_{t+1} from the environment in response to the selected action. Furthermore, the current state s_t evolves to the next state s_{t+1}. The fuzzy model also chooses an action based on the input (state). The collaborative intelligence module chooses one model at each time-step for action selection. Since the RL model is dynamically trained based on feedback from the environment, the environment's responses to the actions selected by the fuzzy model enhance the RL model's learning.

As illustrated in Figure 7.2, this research employs the ϵ-greedy algorithm as the collaborative intelligence module. Herein, a parameter N, randomly sampled from a uniform distribution in the range of [0, 1], is compared with a parameter ϵ. If $\epsilon >$ N, the fuzzy model selects the next action a_t in state s_t. Otherwise, the RL model makes a decision for action selection. Using this approach, the fuzzy model is employed during exploration, allowing the environment to be logically discovered rather than randomly explored, as is typical in the conventional ϵ-greedy algorithm.

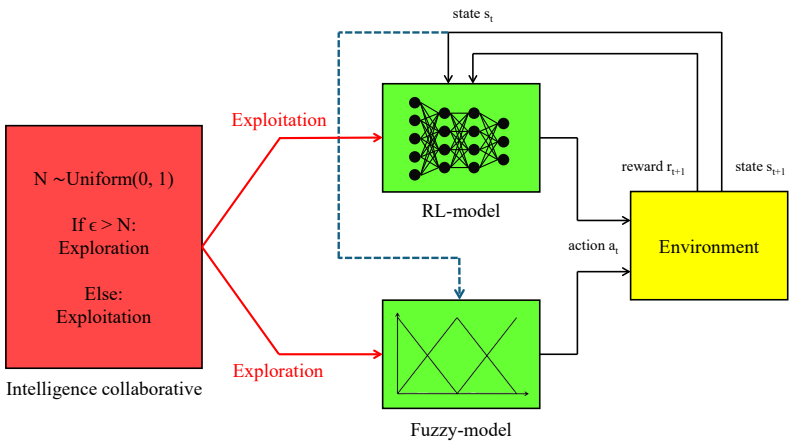


Figure 7.2 The proposed hybrid fuzzy-RL model

Algorithm 6 illustrates the interaction in more detail. It starts by setting an initial value for ϵ, typically between 0 and 1. While the fuzzy model is logically formulated based on rules determined by experts, the RL model is initialized based on predefined learning parameters, such as its architecture. At each time-step, a decision is taken on whether to use the RL model or the fuzzy model. At the end of each episode, the value of ϵ is decayed by a constant factor, less than 1, such 0.995 or 0.9995. This decay gradually reduces the rate of using the fuzzy model as the learning progresses. The idea is to leverage the fuzzy model in the early stages while allowing the RL model to progressively learn from the accumulated experience

gathered from the fuzzy model.

Algorithm 6 The Fuzzy-RL model

```
Initialize an ϵ value
Initialize the fuzzy and RL models
while (1) do                                          ▷ A loop through episodes
    while (1) do                                      ▷ A loop through time-steps
        N ∼ Uniform(0, 1)
        if ϵ > N then
            a_t = Fuzzy model (s_t)                   ▷ The fuzzy model selects an action a_t
        else if ϵ ≤ N then
            a_t = RL model (s_t)                      ▷ The RL model selects an action a_t
        end if
    end while
    ϵ = ϵ *decay factor
end while
```

This study presents a multi-agent framework for allocating tasks to a human and a cobot in a disassembly process as shown in Figure 7.3. This framework is consist of a human agent and a cobot agent, in which each agent is a fuzzy-RL model shown in Figure 7.2. While the human agent selects tasks for the human operator, the cobot agent assigns tasks to the cobotic operator. The agents interact with the environment through Markov decision processes (MDP), which consist of a tuple containing state, action, reward, and next state. The following sections explain the constituent parts of the agents and these interactions in detail.

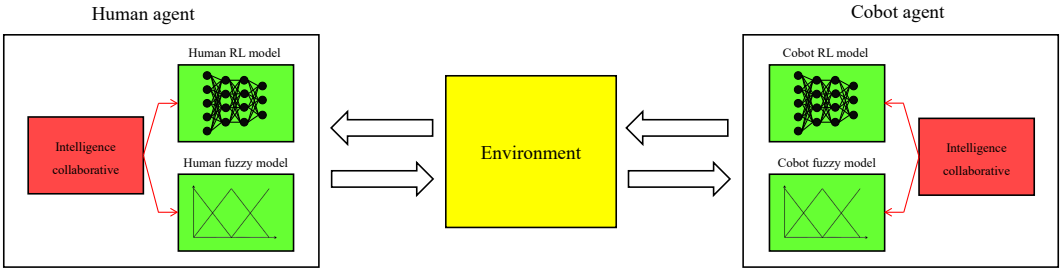


Figure 7.3 The multi-agent fuzzy-RL framework

7.5.2 RL models

This research represents each RL model in the proposed framework with a DQN [124]. It works based on the Q-learning theory and approximates Q-values of actions in each state by using neural networks. By interacting with the environment, it aims to achieve the optimal

Q-function. It estimates the Q-values in such a way that selecting the action with the highest Q-value under policy π maximizes the accumulated reward value, as shown in Equation 7.8. The γ is also a discount factor for incorporating long-term reward values into the immediate reward value. In addition, DQN consists of prediction and target sub-neural networks, which estimate Q values for the current state (Q(s,a)) and Q values for the next states (Q(s′,a′)), respectively. As Equation 7.9 shows, these two values are finally combined with a learning rate of α to form the final Q values. The following presents a detailed explanation of RL models: state, action, reward.

$$\mathrm{Q}^*(\mathrm{s},\mathrm{a}) = \max_{\pi} \mathrm{E}\left[\mathrm{R_t} + \gamma \mathrm{R_{t+1}} + \gamma^2 \mathrm{R_{t+2}} + \ldots \mid \mathrm{s_t} = \mathrm{s}, \mathrm{a_t} = \mathrm{a}, \pi\right] \tag{7.8}$$

$$\mathrm{Q}(\mathrm{s},\mathrm{a}) \leftarrow \mathrm{Q}(\mathrm{s},\mathrm{a}) + \alpha[\,\mathrm{R}(\mathrm{s},\mathrm{a}) + \gamma \arg\max_{\mathrm{a}'} \mathrm{Q}(\mathrm{s}',\mathrm{a}') - \mathrm{Q}(\mathrm{s},\mathrm{a})\,] \tag{7.9}$$

State This research defines the state vector of each model based on the characteristics of product components and the corresponding operator. Figure 7.4 illustrates the state vectors for human and cobot RL models, where $\mathrm{D_i}$, $\mathrm{T_i}$, $\mathrm{H_i}$, and $\mathrm{E_i}$ represent the difficulty, technical feasibility, human safety, and ergonomic risk associated with task i, respectively. The state vector for the human RL model includes the values of difficulty and technical feasibility associated with each task. It also includes the level of ergonomic risk for disassembly of each component. As a safety factor, this vector contains a binary value for each task regarding whether a task requires exposure to hazardous materials or not.

We calculate a difficulty score for each task using the approach presented in [140], which obtains the difficulty score based on the weighted average of several variables. These variables are the size and weight of each component, the force required for its disassembly, a binary variable representing the destructiveness of the operation, the type and characteristics of each liaison, the number of connections of each component with other components, accessibility, and positioning. This study assigns a particular value for each liaison type and a number between 1 and 2 representing the characteristics of connections. While positioning is a preprocessing phase by placing a tool on a connection before starting the operation, accessibility refers to the ease of reaching an area related to the task by hand or tool [80, 82]. This research also considers the feasibility of an operation based on factors such as requiring a special tool or machine, as well as greater strength. On the other hand, the state vector of the cobot agent includes difficulty and technical feasibility scores for each task. Notably, after completing each task by an operator, its corresponding values in the respective state vector will be set to zero.

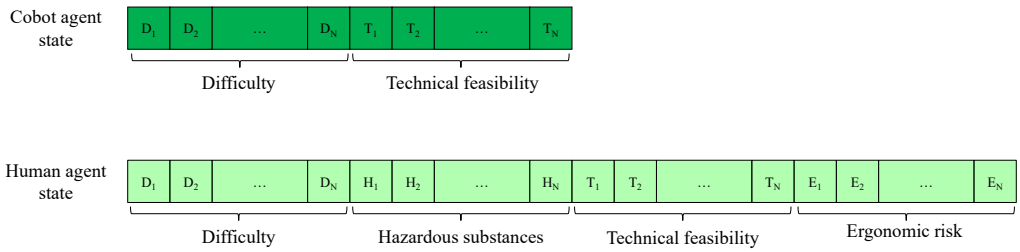


Figure 7.4 The state vectors of the human and cobot agents

Action Since the planning model aims to select tasks for operators, we define the next disassembly task as the action in both agents. Hence, the action space encompasses all available tasks.

Reward As environmental objectives, we aim to minimize the amount of energy consumed by cobot and maximize the quality of recovered components. Moreover, minimizing operation time and tool-changing frequency are considered economic objectives. On the other hand, we consider hazardous materials and ergonomic risks as social factors. In this regard, the model aims to allocate tasks in a way that minimizes human operators' exposure to high-risk and unsafe tasks. To meet the mentioned objectives, we developed a particular reward function for each RL model.

Due to the recursive nature of RL, an agent targets making decisions to maximize cumulative reward values, ensuring the process moves toward reaching sustainable objectives. Equations 7.10 and 7.11 show the reward functions of the human and cobot RL models, which are linear combinations of elements related to objectives. The term R_{human}, the reward function of the human RL model, includes elements related to operation time (R_t), tool change frequency (R_{Tool}), human safety (R_H), ergonomic risks (R_{ergo}), quality of recovered parts (R_q), and a penalty term R_p. In addition, the term R_{cobot}, the reward function of the cobot RL model, comprises elements corresponding to operation time, consumed energy (R_e), recovered parts quality, and a penalty term. The penalty terms in both models target to incentivize the human RL model to choose more difficult and complex tasks and the cobot RL model to select more simple tasks.

Since the cobotic operator can effectively perform high-risk and dangerous tasks, its corresponding reward function does not include social elements. We also assume that the cobotic operator does not need to change tools during the process. Moreover, we cannot observe actual energy consumption as we test the proposed framework in a simulated environment. Hence, it is assumed that R_e rises by one unit for each task carried out by the cobot. On

the other hand, the human RL model's reward function lacks the element corresponding to cobot energy consumption. In addition, each coefficient represents the impact of its associated objective on the reward values. Hence, we can change the influence of an objective in decision-making by adjusting its coefficient. As the objectives are expressed in distinct units, the coefficients also standardize them to ensure uniformity in summation.

$$R_{human} = \alpha_q \cdot R_q - \alpha_t \cdot R_t - \alpha_{ergo} \cdot R_{ergo} - \alpha_H \cdot R_H - \alpha_{Tool} \cdot R_{Tool} - \alpha_p \cdot R_p \tag{7.10}$$

$$R_{cobot} = \alpha_q \cdot R_q - \alpha_t \cdot R_t - \alpha_e \cdot R_e - \alpha_p \cdot R_p \tag{7.11}$$

Graph representation approach This research represents the disassembly process using a graph-based approach introduced in [140]. As shown in Figure 7.5, this approach regards each component as a node and each connection within two components as an edge. In addition, it groups nodes to different levels based on their distances to the origin node. The process starts from the outermost level. After disassembling each node, other nodes in this level or nodes connected to the disassembled nodes can be selected for disassembly. The rest of the nodes cannot be selected due to precedence relations. In order to integrate this algorithm into the RL framework, a vector with a length equal to the number of nodes is added to the output of the DQN network (actions' Q-values) in each time-step. This vector contains zero values for eligible nodes to be disassembled and -100 for the rest of them. This strategy ensures nodes with precedence dependencies will not be selected by the RL models.

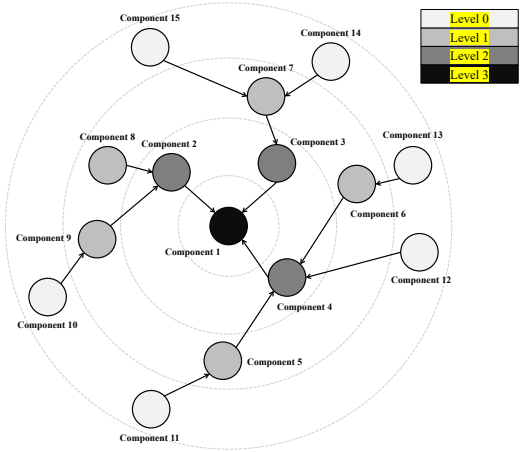


Figure 7.5 The graph representation approach

7.5.3 Fuzzy models

A fuzzy inference system (FIS) consists of several modules: a fuzzifier, an inference engine, and a defuzzifier. First, the input crisp value is fuzzified. Next, the inference module makes decisions based on predefined rules. Finally, the fuzzy output sets are defuzzified. In this part, we explain the developed FIS modules in detail.

We define 'Effectiveness' as a fuzzy number, which serves as the output of FIS. This value represents the appropriateness of each task for the corresponding operator. The inputs to this system are the difficulty and technical feasibility of a task. The FIS evaluates how effective the task is for the given operator according to its degrees of difficulty and feasibility. Each time a fuzzy model is called, the available nodes for disassembly are provided to the FIS. After assessing the effectiveness of each node, the node with the highest effectiveness is selected. To represent 'Effectiveness', we use a triangular fuzzy number, as shown in Figure 7.6, which categories this value into three classes: low, medium, and high.

On the other hand, in the fuzzification step, we represent the difficulty and technical feasibility values, which serve as inputs to the FIS, by using fuzzy triangular numbers. Furthermore, the centroid defuzzification method, shown in Equation 7.12, converts the fuzzy output set to a single crisp value. In this case, x and $\mu(x)$ determine the output fuzzy set and the membership function, respectively. In addition, z represents the final output crisp value.

$$z = \frac{\int x \cdot \mu(x) dx}{\int \mu(x) dx} \tag{7.12}$$

In this problem, there are two agents: a human-agent and a cobot-agent. As mentioned in Section 7.5.1, each agent consists of a fuzzy-RL model, illustrated in Figure 7.2. Section 7.5.2 defines two RL models, one for each agent. Correspondingly, we define two fuzzy models: a human fuzzy and a cobot fuzzy. Each fuzzy model is a FIS that assigns tasks to its respective operator. Tables 7.2 and 7.3 illustrate the rules of the fuzzy models, defined by expert knowledge. As shown in Table 7.2, the cobot achieves High or Medium effectiveness when its tasks are of Low difficulty and feasibility levels are High or Medium. However, the effectiveness value decreases to Low when the difficulty level is High. For tasks with Medium difficulty, the effectiveness value depends on the feasibility level that falls into Low or Medium. In Table 7.3, an effectiveness value of Low corresponds to a feasibility value of Low. For other cases, the effectiveness value is set to Medium or High depending on the difficulty values. Notably, the value of Low is assigned to effectiveness in the case of Low difficulty and High feasibility tasks, which logically aligns with the simpler nature of tasks typically performed by cobots.

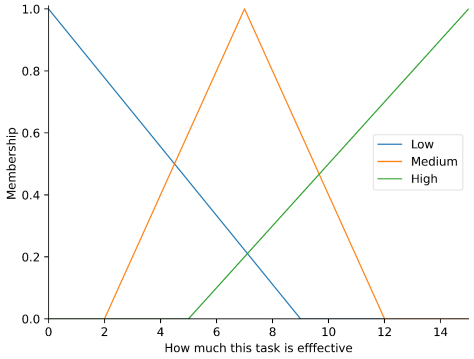


Figure 7.6 Effectiveness fuzzy membership

Table 7.2 Cobot fuzzy rules

Number of rule	Difficulty level	Cobot Technical feasibility	Effectiveness
1	Low	High	High
2	Low	Medium	Medium
3	Low	Low	Low
4	Medium	High	Medium
5	Medium	Medium	Medium
6	Medium	Low	Low
7	High	High	Low
8	High	Medium	Low
9	High	Low	Low

Table 7.3 Human fuzzy rules

Number of rule	Difficulty level	Human Technical feasibility	Effectiveness
1	Low	High	Low
2	Low	Medium	Medium
3	Low	Low	Low
4	Medium	High	Medium
5	Medium	Medium	Medium
6	Medium	Low	Low
7	High	High	High
8	High	Medium	High
9	High	Low	Low

7.5.4 Environment modeling

As explained in Section 7.5.1, the developed fuzzy-RL approach dynamically learns by interacting with the environment. This research mathematically represents the environment and the interactions in order to provide a relevant output for a selected action (task performed by an operator). While the proposed approach learns from this environment, its interactive and dynamic nature enables it to perform in real-world scenarios and make decisions under uncertain conditions.

Tables 7.4 and 7.5 show the environment representation for the cobot and human operators, respectively. By considering the difficulty and technical feasibility scores for a task, the environment provides operation time, recovered quality, and penalty. This study uses four threshold values —T_{D_1}, T_{D_2}, T_{F_1}, and T_{F_2}— to distinguish the boundaries between Low, Medium, and High levels of difficulty and technical feasibility scores. These thresholds are determined by experts based on historical data and their professional expertise in evaluating task complexity in HRC disassembly projects. Experts assessed the performance of operations to determine the points at which the success rate of an operation decreases significantly due to its extreme complexity. If a task achieves a low technical feasibility score, it is considered as a failure. The values of these thresholds are presented in Section 7.6.

When a task is successfully completed, the reward value is calculated using Equations 7.10 and 7.11, depending on whether the operator is a cobot or a human. If a task cannot be performed, the environment assigns a reward value of -10 to penalize the respective agent.

Table 7.4 Time, recovered quality, and penalty in various conditions for the cobot operator

Technical feasibility	**Difficulty**		
	Low	**Medium**	**High**
Low	time = 3 reward = -10 'Task not completed'	time = 3 reward = -10 'Task not completed'	time = 3 reward = -10 'Task not completed'
Medium	time = 2 quality = 7 penalty = 0 'Task completed'	time = 3 quality = 2 penalty = 0 'Task completed'	time = 3 reward = -10 'Task completed'
High	time = 1 quality = 10 penalty = 0 'Task completed'	time = 2 quality = 7 penalty = 0 'Task completed'	time = 3 reward = -10 'Task completed'

7.5.5 Failure operations probability

In a disassembly process, a cobot may fail to complete some tasks because of abnormal conditions, including machine failure, wear and tear tools, cobot inefficiency, or any other external interruption, which direct the task sequence planning process. In such cases, the

Table 7.5 Time, recovered quality, and penalty in various conditions for the human operator

Technical feasibility	**Difficulty**		
	Low	**Medium**	**High**
Low	time = 3 reward = -10 'Task not completed'	time = 3 reward = -10 'Task not completed'	time = 3 reward = -10 'Task not completed'
Medium	time = 1 quality = 10 penalty = -10 'Task completed'	time = 1 quality = 10 penalty = 0 'Task completed'	time = 3 reward = -10 'Task completed'
High	time = 1 quality = 10 penalty = -10 'Task completed'	time = 1 quality = 10 penalty = 0 'Task completed'	time = 3 reward = -10 'Task completed'

model should reconfigure the task sequence from its original outline. Therefore, it is crucial to consider task failure in the modeling process to ensure the model's effectiveness in industrial settings.

This research study models task failure probabilities by a Beta distribution, determined over the interval of [0,1]. This distribution delivers great flexibility to model various shapes. Furthermore, the nature of this distribution, which is over [0,1], fully corresponds with the task failure probability, falling between 0 and 1. As a result, this distribution is suitable for representing the inherent variability and uncertainty in failure probability during disassembly processes.

Equations 7.13 and 7.14 show the distribution, in which $B(\alpha, \beta)$ and $f(x; \alpha, \beta)$ represent the Beta function and the probabilistic distribution function, in that order. In addition, Equation 7.15 denotes the failure probability (P_{failure}) for a given task i, where D[i] and TF[i] refer to the difficulty and technical feasibility scores associated with task i. The parameters alpha and beta directly affect the failure probability by determining the shape of the distribution. Increasing α drives the probability of values closer to 1 (presenting a higher failure probability), whereas higher β shifts the probability of values closer to 0 (indicating a lower failure probability). Moreover, $\alpha = \beta = 1$ results in a uniform distribution. These two parameters can be adjusted based on historical data or by experts. In addition, it is assumed that difficulty values directly, and technical feasibility values inversely, impact the failure probability. Notably, if this failure probability exceeds a certain threshold, the corresponding operation fails.

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt \tag{7.13}$$

$$f(x;\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \tag{7.14}$$

$$\mathrm{P}_{\text{failure}} = f(x;a,\beta) * \mathrm{D[i]} * \frac{1}{\mathrm{TF[i]}} \tag{7.15}$$

7.5.6 The multi-agent fuzzy-RL algorithm for HRC disassembly planning

In this part, we provide a brief overview of the proposed algorithm for HRC disassembly planning. Figure 7.7 illustrates the algorithm's flowchart, which starts with configuring the agents' settings. Herein, the fuzzy models are configured based on the rules discussed in Section 7.5.3, while the RL models' parameters are initialized randomly. Next, human and cobot agents select actions in order. The algorithm then proceeds along a loop, consisting of two parallel paths, each for one agent. Subsequently, a condition block checks the completion of tasks in each path. An agent waits if the corresponding task has not yet been completed. Otherwise, it receives a reward for the performed action and evolves to the next state. Then, it stores the corresponding MDP (state, action, reward, next state) in the replay memory. A batch of data is sampled from the memory to train the respective agent. In the next step, each agent's state vector is updated based on the actions completed by another agent. This prevents each agent from selecting an action already completed by the other. Following that, a condition block selects one of the corresponding RL or fuzzy models for decision-making by comparing the ϵ value to a value derived from a normal distribution. This condition block plays the role of the collaborative intelligence module. The process will be terminated once all tasks are completed. Otherwise, the algorithm will reduce the ϵ value by a constant factor and repeat the loop.

7.6 Results and discussion

This section evaluates and validates the developed framework under different scenarios. First, it explains the dataset applied to evaluate the model, the evaluation protocol, and parameter settings. Next, it discusses different sensitivity analyses, including trade-offs in the exploitation-exploration dilemma, real-time decision-making, operation failure, and trade-offs in sustainable objective importance. Lastly, it presents a comprehensive discussion concerning all experiments.

Evaluation protocol, data, and parameter settings In order to evaluate the performance of the proposed models, we calculate the sum of reward values in all time steps of an

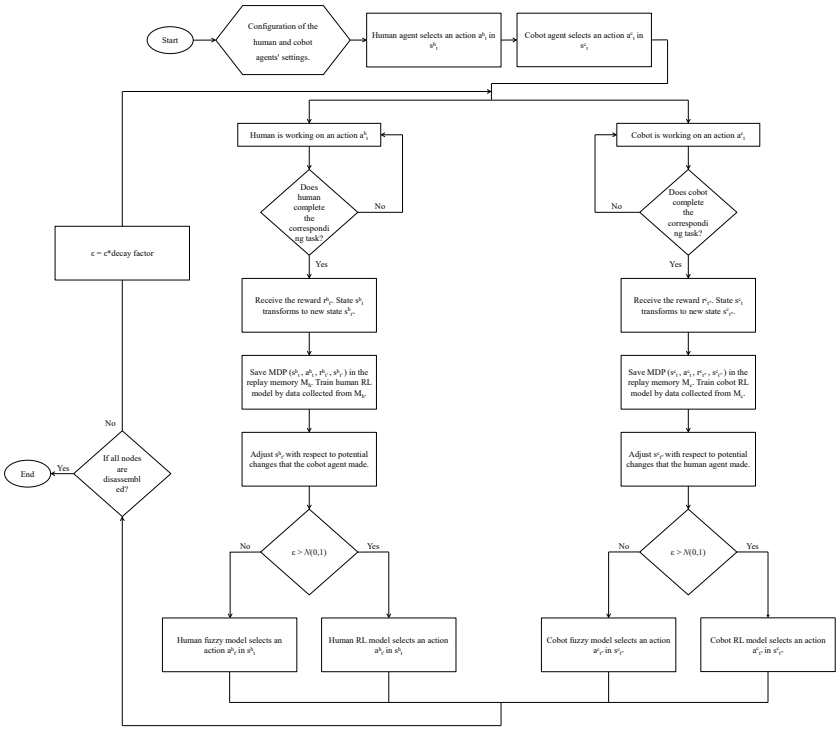


Figure 7.7 The flowchart of the proposed HRC disassembly planning algorithm

episode, which is shown in Equation 7.16. In this case, r_t and T indicate the reward value in time step t and the number of time steps in an episode, respectively. Accordingly, we evaluate the model performance by analyzing and comparing the cumulative reward values in episodes. Since this study presents a multi-agent framework, the evaluation is based on the summation of reward values of both agents.

While comparison with other optimization methods, such as heuristic algorithms, can deliver more understanding, this research focuses on evaluating the performance of the novel fuzzy-RL framework against the classical RL model. We believe this approach sufficiently proves the robustness and effectiveness of the proposed method.

This research employs an electric meter presented in [154] as a case study to validate the model. Table 7.7 provides the features associated with each component. Since the referenced article does not provide all the required features for the product components, we consider other attributes, along with the attributes mentioned in the referenced article. In this research, we set the γ and α values to 0.1 and 0.2 respectively. The replay memory size is configured to 100, and the target sub-network is updated every 50 iterations. We also set T_{D_1}, T_{D_2}, T_{F_1}, and T_{F_2} as 14.9, 19, 1, and 5, respectively. Furthermore, during the sensitivity analysis, we alter one parameter while keeping the values of other parameters constant.

$$R = \sum_{t=1}^{T} r_t \tag{7.16}$$

Sensitivity analysis exploitation-exploration This part aims to provide a trade-off analysis between the exploration/exploitation dilemma by applying the developed model to the dataset under different configurations. These configurations are defined by varying four parameters, ϵ, decay factor, number of neural network layers, and exploration type, which are shown in Table 7.6. The parameters ϵ and decay factor directly affect the number of explorations/exploitations of the model during the learning process. Moreover, a neural network's number of layers typically impacts its generalizability and ability to learn complex patterns. We configure the number of layers to three and four and set the parameter ϵ to three values: 0.1, 0.01, and 0. In addition, we adjust the decay factor to 0.975, 0.9975, and 0.99975. We also configure the model in two random and fuzzy exploration settings. In the fuzzy exploration setting, we employ the fuzzy-RL framework shown in Figure 7.3. In contrast, only RL models are used in the random exploration setting, and actions are selected randomly in the exploration mode. We selected six models that outperform others due to their high and stable reward values. Figure 7.8a-7.8f illustrate the reward values of these models, which are set based on Configurations 1, 5, 9, 12, 18, and 21. Plots associated with Configurations 5, 9, and 21 exhibit more fluctuations and lack stability. However, the plot shown in Figure 7.8d achieves high reward values around episode 150 but continues to significantly fluctuate. The plots shown in Figure 7.8a and 7.8e reach lower reward values than the plot shown in Figure 7.8d, but they converge faster and exhibit fewer fluctuations. Furthermore, Figure 7.8e, corresponding to Configuration 18, demonstrates greater stability compared to another plot.

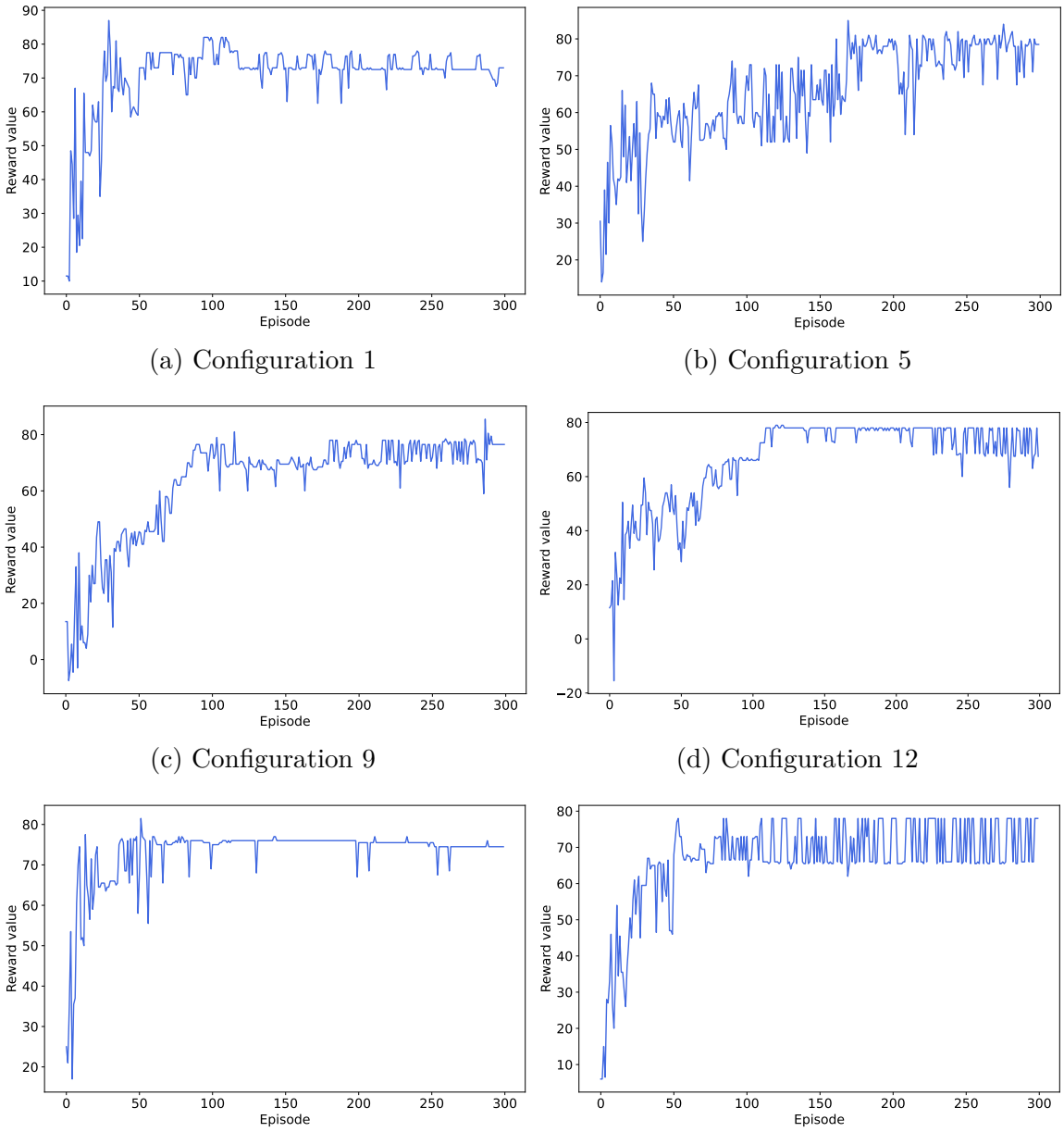
Increasing reward values even slightly in regions where the model has converged is remarkably challenging. The model under Configuration 12 delivers higher reward values in these regions than in other configurations. However, this setting provides much fluctuation, resulting in poor stability. Assuring stability in this field is crucial, as an unstable model ineffectively allocates tasks, significantly decreasing the profitability and productivity of the process. In other words, Configuration 12 does not provide sustained performance despite boasting higher reward values. On the other hand, Configurations 1 and 18 provide enhanced stability compared to Configuration 12. These settings correspond to exploitation without exploration (a=0) and fuzzy exploration, respectively. Furthermore, Configuration 18 provides more stable performance compared to Configuration 1, proving the effectiveness of the fuzzy exploration approach.

Table 7.6 The model configurations based on the impactful factors in the exploitation/exploration trade-off

Configuration	Number of layers	ϵ	decay factor	Exploration
1	3	0	-	-
2	4	0	-	-
3	3	0.01	0.975	Fuzzy
4	3	0.01	0.9975	Fuzzy
5	3	0.01	0.99975	Fuzzy
6	3	0.1	0.975	Fuzzy
7	3	0.1	0.9975	Fuzzy
8	3	0.1	0.99975	Fuzzy
9	3	0.01	0.975	Random
10	3	0.01	0.9975	Random
11	3	0.01	0.99975	Random
12	3	0.1	0.975	Random
13	3	0.1	0.9975	Random
14	3	0.1	0.99975	Random
15	4	0.01	0.975	Fuzzy
16	4	0.01	0.9975	Fuzzy
17	4	0.01	0.99975	Fuzzy
18	4	0.1	0.975	Fuzzy
19	4	0.1	0.9975	Fuzzy
20	4	0.1	0.99975	Fuzzy
21	4	0.01	0.975	Random
22	4	0.01	0.9975	Random
23	4	0.01	0.99975	Random
24	4	0.1	0.975	Random
25	4	0.1	0.9975	Random
26	4	0.1	0.99975	Random

Sensitivity analysis on real-time decision-making As mentioned in Section 7.5, various factors may deviate a process from the ideal path. Thus, a disassembly planning model should be able to cope with these uncertainties by making real-time decisions. This sensitivity analysis aims to evaluate the real-time decision-making capability of the proposed approach. In this way, the execution time determined for each task is not constant and may vary. Equation 7.17 represents the execution time considered for the i-th task, in which T is a predefined fixed number and t is a variable number, which is randomly selected based on a uniform discrete distribution. Accordingly, the Time vector with length 50 including the numbers 0, 1, 2, and 3 is defined as shown in Equation 7.18. Equation 7.19 denotes the probability of selecting a number X from the Time vector, given that x_i represents the i-th unique value (0,1,2, or 3) in the vector, and the function δ, shown in Equation 7.20, returns 1 if the inputs are the same. In addition, N indicates the vector length. In this way, $P(X = 0) = 0.94$, and $P(X = 1) = P(X = 2) = P(X = 3) = 0.002$.

$$T^{'} = T + t \tag{7.17}$$



(a) Configuration 1

(b) Configuration 5

(c) Configuration 9

(d) Configuration 12

(e) Configuration 18

(f) Configuration 21

Figure 7.8 Reward values resulted from different configurations

$$\text{Time} = [\underbrace{0, 0, 0, 0, \ldots, 0}_{47\text{numbers}}, 1, 2, 3] \tag{7.18}$$

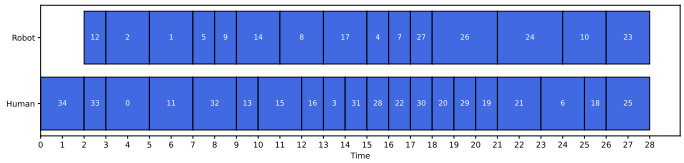
$$t = P\left(X = x_i\right) = \frac{1}{N} \sum_{j=1}^{N} \delta\left(x_i, x_j\right) \tag{7.19}$$

$$\delta\left(\mathrm{x_i} \cdot \mathrm{x_j}\right) = 1 \text{ if } \mathrm{x_i} = \mathrm{x_j} \text{ Otherwise } 0 \tag{7.20}$$

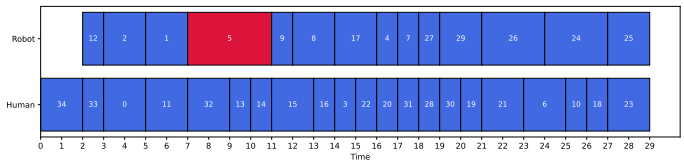
Figures 7.9a-7.9c illustrate the task sequences generated in scenarios with variable time settings. Blue blocks indicate tasks with predetermined disassembly times (t=0 in Equation 7.17), while red blocks represent tasks with unpredicted times (t≠0 in Equation 7.17). Thus, all the tasks have taken their expected time in task sequence 1. In contrast, tasks 5 in sequence 2, as well as tasks 33 and 18 in sequence 3, exceeded their expected times. As shown in Figures 7.9a-7.9c, the model adopts a consistent pattern for task allocation before encountering the tasks with unexpected times (red blocks). The model dynamically generates different task sequences as completing a task requires longer times (the process change from its ideal flow). For example, in task sequences 1 and 3, the model allocated task 14 to the cobot after completing tasks 5 and 9. However, in task sequence 2, the model assigned this task to the human operator as the cobot completed task 5 slower than expected. Rather than using the pre-planned sequence, the model allocates alternative tasks to the operators in real time to optimize the process based on new conditions.

Sensitivity analysis on sustainable objectives importance This sensitive analysis compares model performance under different sustainable objective configurations. It investigates the importance of social factors (ergonomic risk and human safety) in the generated task sequences. Figure 7.10 illustrates the task sequences associated with $\alpha_{\mathrm{H}} = \alpha_{\mathrm{ergo}} = 0.1$, 10, and 50 in episode 300, in which alpha α_{H} and alpha α_{ergo} are coefficients for the reward terms related to human safety and ergonomic risk in Equation 7.10. Herein, the color of each block is associated with its risk level, where the red and orange blocks present high and medium-risk tasks, respectively. In addition, the green blocks indicate risk-free tasks for the human operator. As shown in Figure 7.10a-7.10c, with increasing α_{ergo} and α_{H}, the number of risky operations included in the task sequence of the human agent decreases. In this case, the sequence generated for the human operator with $\alpha_{\mathrm{H}} = \alpha_{\mathrm{ergo}} = 0.1$ involves more risky tasks compared to other setups. By adjusting the coefficients of the objectives in the reward functions, the corresponding task sequences change. This allows us to customize the disassembly process to prioritize sustainable objectives according to their importance.

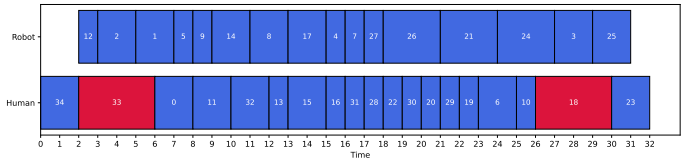
To support practical applications and enhance user interaction with the proposed framework, we developed a GUI, shown in Fig. 7.11, using Python's Streamlite library. As illustrated in Fig. 7.11a, this interface initially requests users to input three values between 0 and 100, representing the importance of desired economic, social, and environmental objectives. These inputs are then used as weights for respective objectives in the reward functions. This feature



(a) Task sequence 1

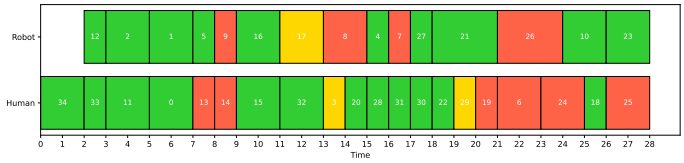


(b) Task sequence 2

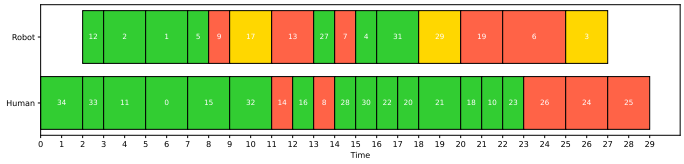


(c) Task sequence 3

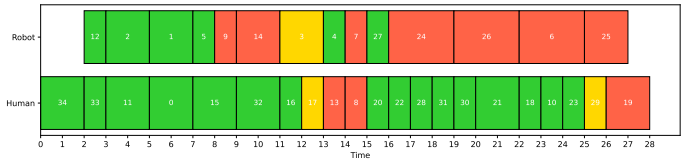
Figure 7.9 Sensitivity analysis on real-time decision-making



(a) $\alpha_{social} = 0.1$



(b) $\alpha_{social} = 10$



(c) $\alpha_{social} = 50$

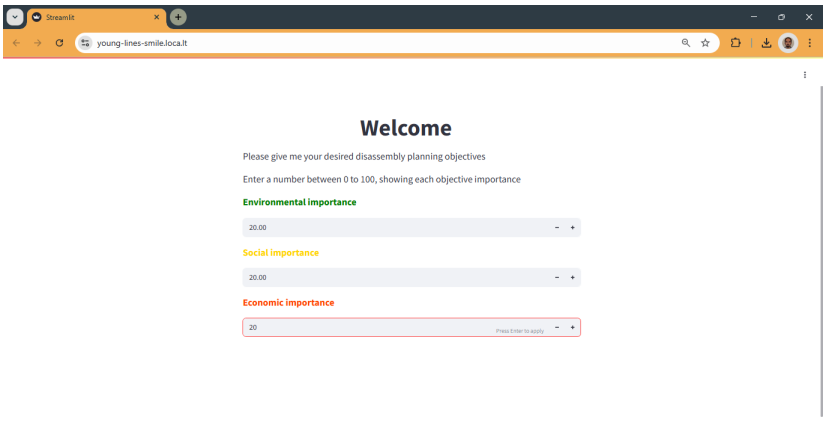
Figure 7.10 Sensitivity analysis on sustainable objectives importance

enables users to customize disassembly processes based on their specific requirements. For instance, in an industrial setting in which environmental concerns are highly prioritized, manufacturers can fix the environmental input at 100 and assign lower values to other objectives. The GUI provides a clear visualization of inputs and results to facilitate decision-making. For example, as shown in Fig. 7.11b, a donut chart shows the relative importance of the sustainability objectives, enabling users to immediately validate their input visually. Following this, the learning process is launched, and the corresponding task sequences are displayed in an intuitive format, allowing users to assess the generated disassembly plan. By combining input customization and result visualization, the GUI bridges the gap between theoretical models and industrial usage. It ensures that users can interact with the framework to optimize processes according to their requirements.

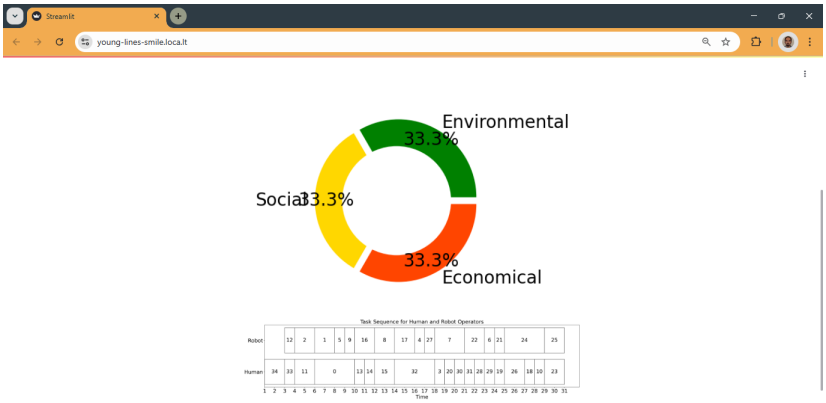
Sensitivity analysis on failure operations As discussed in Section 7.5.4, this study involves the failure of an operation by a cobotic operator, leading the process out of its expected sequence. This sensitivity analysis aims to evaluate the model's capability to cope with these failing scenarios. We set the values of parameters α and β in Equations 7.13-7.15 as 0.2 and 4, respectively, and integrated the failure probability into the planning process. Figure 7.12a-7.12c show the three task sequences, in which green blocks represent tasks successfully completed. Red blocks indicate tasks that the cobot failed to complete, whereas yellow blocks display tasks that were reallocated to the human operator after failing by the cobot. In this way, in sequence 1, all tasks were effectively accomplished. However, tasks 1, 5, and 21 in sequence 2, along with tasks 1 and 23 in sequence 3 have failed. As shown in Figure 7.12b and 7.12c, these tasks were subsequently assigned to the human operator. Furthermore, it is evident that the process takes longer compared to normal conditions (task sequence 1).

7.7 Conclusion

In recent years, the use of cobots has become increasingly popular in industrial disassembly processes. Cobots efficiently handle repetitive and simple tasks that human operators may struggle to perform accurately due to fatigue or distraction. Furthermore, cobots can complete tasks that expose humans to high ergonomic risks or involve handling toxic materials. However, cobots cannot perform complex and difficult tasks due to their lack of strength and flexibility. Combining human capability in performing complex and challenging tasks with the precision of cobots significantly enhances the productivity and quality of processes. Planning HRC disassembly processes by generating task sequences for operators is a critical

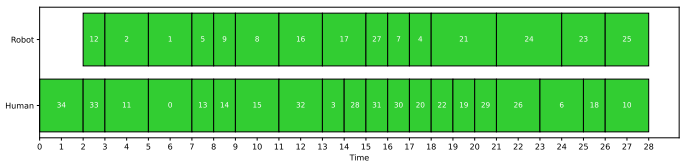


(a) User inputs

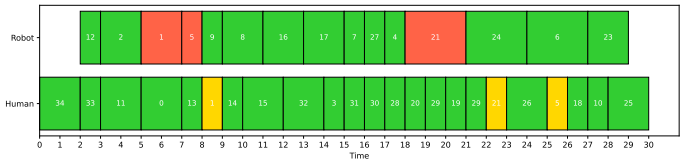


(b) Customized outputs

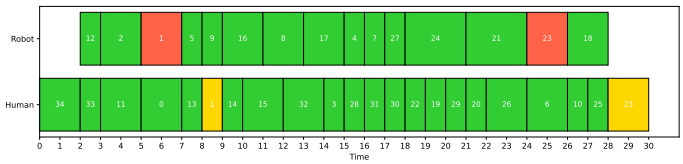
Figure 7.11 The developed GUI



(a) Task sequence 1



(b) Task sequence 2



(c) Task sequence 3

Figure 7.12 Sensitivity analysis on operation failure

step in configuring these processes. A planning approach should optimize task allocation process based on multiple objectives, such as minimizing energy consumption, operation time, and ergonomic risks on the one hand, and maximizing the quality of recovered parts and human safety on the other hand. Additionally, various factors, including human errors, variable product conditions, machine failures, tool wear and tear, or any external disruptions, may deviate the process from the preplanned direction, leading to task prolongation or failure. Therefore, a robust planning model should make real-time decisions based on online conditions to address these uncertainties effectively.

This research introduces a sustainable multi-agent framework based on RL for HRC disassembly planning. Capable of making real-time decisions, the framework dynamically allocates tasks based on online conditions. It consists of two agents—human and cobot—where each agent selects tasks for its respective operator. Each agent is an RL model that leverages a fuzzy-based exploration approach, logically selecting actions during the exploration phase of the ϵ-greedy algorithm. As the fuzzy approach is configured according to experts' knowledge, this study incorporates human expertise within an RL-based structure. In addition, the proposed framework includes two reward functions (human and cobot) that are linear combinations of several sustainable elements. These functions are powerful key performance

indicators in industrial settings. The reward functions emphasize minimizing operation times and tool changes frequency, directly impacting on productivity. The functions also involve environmental elements, such as circularity and energy consumption. Social factors, such as human safety and ergonomic risk, are also integrated into the functions. By presenting an adjustable GUI, manufacturers can set the coefficients associated with these sustainable elements to customize the process according to their unique priorities.

Experimental results from applying the proposed approach to a case study reveal that the framework, when using fuzzy exploration, achieves more stable performance compared to other configurations. Additionally, this article presents multiple sensitivity analyses to validate the model's performance in scenarios with varying task operation times, operation failure probabilities, and varying importance of sustainability objectives.

In future research, we plan to implement the proposed framework in laboratory settings to develop an industrial prototype. The use of the copilot concept and filtering actions based on expert knowledge could significantly improve the performance of an RL model in larger-scale problems, where selecting an action from many possible actions is more challenging and complex. We plan to test the fuzzy-RL model in higher-dimensions problems, such as recommender systems and transportation, to evaluate the performance of a conventional RL model with fuzzy exploration. We also aim to conduct a comparative analysis of the fuzzy-RL approach with other optimization methods, including various heuristic algorithms. Moreover, we aim to explore other methods for incorporating experts' knowledge into an RL model and provide a comparative analysis with these approaches.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [grant number RGPIN-2020-05565].

Data availability statement

The data that support the findings of this research study are available from the corresponding author, with reasonable request.

Appendix

Table 7.7 The case study feature values

Component	Positioning	Accessibility	Weight	Strength	Size	Destructive	Liaisons scores	Liaisons properties	Ergonomic risk	Hazardous substances	Human feasibility	Cobot feasibility	Tool
1	2	1	2.4	4	3.5	1	6	1.4	0	0	13	4	1
2	1.2	1.6	2	1	2	0	2	1.2	0	0	14	4	1
3	1.2	1	2	2	2	0	3	1.3	0	0	3	3	2
4	2	2	2.2	1	4	0	2	1.4	0.5	0	2	2	1
5	1.2	1.6	2.2	2	3.5	0	4	1	0	0	12	11	3
6	5	1	2	1	2	0	2	1.1	0	0	17	15	1
7	1.9	2	4	4	4	1	10	2	0	1	13	4	1
8	1.2	2	2	2	2	0	3	1.2	1	0	14	16	1
9	2	1.6	2.2	2	2	0	2	1.1	0	1	10	5	3
10	2	1	2	1	2	0	4	1.5	0	1	11	17	2
11	1.2	1.6	2.2	1	2	0	2	1.7	0	0	15	3	1
12	1.2	1	4	4	4	1	10	1.8	0	0	14	10	1
13	1.2	2	2	1	2	0	2	1.1	0	0	13	19	3
14	1.2	2	2.2	1	3	0	3	1.9	0	1	15	13	1
15	5	2	2	2	2	0	4	1.5	1	0	16	11	2
16	2	1	2.4	4	3.5	0	6	1.4	0	0	13	4	1
17	1.2	1.6	2	1	2	0	2	1.2	0	0	14	4	1
18	1.2	1	2	2	2	0	3	1.3	0.5	0	3	3	2
19	2	2	2.2	1	4	0	2	1.4	0	0	2	2	1
20	1.2	1.6	2.2	2	3.5	0	4	1	0	1	12	11	3
21	5	1	2	1	2	0	2	1.1	0	0	17	15	1
22	5	1	2.4	4	3.5	0	10	2	0	0	13	4	1
23	1.2	2	2	2	2	0	3	1.2	0	0	14	16	1
24	2	1.6	2.2	2	2	0	2	1.1	0	0	10	5	3
25	1	1	2	1	3	0	4	1.5	0	1	11	17	2
26	1.2	1.6	2.2	1	2	0	2	1.7	0	1	15	3	1
27	1.9	1.9	4	4	4	1	6	1.9	0	1	14	10	1
28	1.2	2	2	1	2	0	2	1.1	0	0	13	19	3
29	1.2	2	2.2	2	2	0	3	1.9	0	0	15	13	1
30	5	2	2	2	2	0	4	1.5	0.5	0	16	11	2
31	1.2	1.6	2	1	2	0	6	1.2	0	0	15	16	1
32	2	1	2.2	4	3.5	0	2	1.1	0	0	14	15	1
33	2	2	4	1	2	0	3	1.5	0	0	13	17	3
34	1.2	1.6	2	2	2	0	2	1.7	0	0	15	13	1
35	1	1	2.2	2	4	0	4	1.8	0	0	16	10	2

CHAPTER 8 CONCLUSION AND GENERAL DISCUSSION

Recently, the use of HRC in disassembly processes has gained popularity and has been significantly explored by many researchers. HRC disassembly promises progress in the industry by taking advantage of the flexibility of human operators and the high precision of cobots. Despite all the benefits of HRC disassembly over traditional manual disassembly, this field still faces serious challenges. This research study aims to address the existing limitations and fill the gaps in the literature by presenting several novel contributions across four articles. The following is a summary of these articles.

8.1 The first article

This study proposes a novel multi-agent RL model based on DQN for disassembly planning in a human-robot collaborative environment, in which the model sequentially assigns tasks to a cobot and a human operator. A key contribution was the real-time adaptation of the model, which dynamically adjusts task allocation based on evolving process conditions. In addition, this work integrates part difficulty scores, calculated based on the physical and geometric characteristics of the components, into the decision-making process to ensure that tasks are assigned based on operator capability. The research also presents a graph-based approach to modeling an EoL product and the relationships between its components. The approach models each component with a node and each connection between two components with an edge. Furthermore, this study considers the skill levels of human operators in decision-making to improve process efficiency through skill-based assignments.

In order to comprehensively and effectively validate the presented model, this research study uses qualitative and quantitative comparative analyses. Here, the most essential evaluation criteria for RL models are the cumulative reward value per episode and the convergence time in reward plots. Furthermore, conducting sensitivity analysis and fine-tuning the impactful parameters of the problem, such as ϵ, γ, and the replay memory size, not only improves the models' performance but also enables us to achieve a more profound understanding of the conditions governing the problem. The analyses show that the proposed model significantly outperforms the baseline models in terms of features involved in the decision-making process and convergence speed. Compared to static optimization methods, this approach offers greater adaptability to uncertainties, such as human skill levels, cobot limitations, and product variability. However, a limitation of this study was that it did not explicitly consider sustainability criteria, such as energy consumption and ergonomic risks, in the optimization

process. This gap was addressed in subsequent research.

8.2 The second article

This study extends the first article by formulating a multi-agent RL model that optimizes HRC disassembly processes with respect to sustainability indicators, as defined in Section 1.2. Unlike conventional optimization methods that rely on fixed task sequences, this model dynamically allocates tasks by balancing economic, environmental, and social considerations. This study introduces sustainability-aware reward functions that incorporate energy consumption, quality of recovered parts (environmental objectives), ergonomic risk and human safety (social objectives), along with operation time and tool change frequency (economic objectives). The study also uses fuzzy logic to model uncertain factors in the environment, such as degrees of difficulty and technical feasibility. The use of fuzzy logic in RL allows the model to cope better with the possibilistic nature of uncertainties in the disassembly process, making the model's training simulation more reflective of the real-world conditions.

The study investigates different RL configurations and compares performance metrics, such as learning efficiency, convergence stability, and adaptability to uncertainties. Experimental results demonstrate that the proposed model not only maintains a high level of economic benefits in task allocation but also meets environmental and social considerations.

Note that this study focuses on operational sustainability indicators and does not address potential second-order or rebound effects (e.g., social impacts of job displacement). Furthermore, although other operational indicators, such as life cycle assessment (LCA), are available, exploring these indicators falls beyond the scope of this study.

8.3 The third article

This study introduces a new approach to assessing ergonomic risk in disassembly processes by integrating real-time video processing, RULA, REBA, and fuzzy logic. Unlike traditional ergonomic assessments, which often fail to capture the dynamic nature of human movements, the proposed method dynamically assesses ergonomic risks by using a multi-camera vision system to capture human movements in real time. First, the proposed approach captures frames from three cameras placed in the disassembly cell. It then detects the wrist, elbow, shoulder, and hip joints on both sides. Following that, the angles between the different joints are calculated.

For ergonomic risk assessment, the proposed approach first evaluates the posture risks of

shoulders and elbows in each task based on the angles from two directions—side and top views— by considering the scoring methodology of REBA and RULA. Then, it integrates the force load of each task using MVC with the posture risks to calculate the final risk for each task. Subsequently, it computes the cumulative risk of each body part based on the task duration compared to the cycle time. Finally, the total ergonomic risk of each body part is evaluated using five fuzzy rules. We validated the proposed approach by experimenting in a laboratory setting, ensuring its effectiveness in extracting joint angles, classifying body posture risks, and assessing cumulative ergonomic risks.

8.4 The fourth article

This paper introduces a fuzzy-guided exploration approach within an RL model to enhance disassembly planning processes. In this approach, a fuzzy model adjusts the exploration phase based on environmental conditions. Given the poor performance of RL models in initial iterations, the fuzzy model makes decisions in the early time-steps, offering valuable feedback from the environment to train the RL model. As learning progresses, the fuzzy model is gradually used less and less until the RL model optimizes the process independently. Since the fuzzy model is logically configured based on experts' knowledge, this research combines an expert-driven approach with a learning-driven method, resulting in a robust hybrid solution. This new approach allows RL agents to make more informed decisions in uncertain scenarios and improves the overall robustness of the planning process.

This study conducts several new sensitivity analyses to investigate the performance and validate the model under different conditions. This study presents a trade-off analysis on exploitation-exploration. It also provides a trade-off analysis among the sustainability objectives (Section 1.2) in the decision-making process. It compares the generated task sequences by adjusting the importance of different objectives. Furthermore, a sensitivity analysis is conducted to evaluate the model's performance under uncertain conditions with variable execution times for each task. In addition, this research presents another sensitivity analysis by introducing a failure probability for tasks performed by cobots, aiming to assess the model's ability to reconfigure task sequences in response to dynamic conditions. This research also develops a GUI that enables users to customize the disassembly process according to their individual requirements by adjusting the sustainable objectives. Experimental validation using an electronic board as a case study shows that the fuzzy-guided RL approach achieves more stable rewards compared to conventional RL configurations.

8.5 Implications

As academic implications, this thesis delivers multiple novelties by introducing three new multi-agent RL models for HRC disassembly planning. These models aim to cover the weak points of the previous works by offering comprehensive and sustainable planning approaches that optimize the process based on dynamic conditions by incorporating many factors, including components' characteristics, operator skill level, and execution time. Additionally, this research proposes a graph-based method for representing a product's components and their dependencies. It also introduces a fuzzy approach to modeling uncertain parameters in the environment. In order to improve the performance of RL-based disassembly planning models, this research presents a novel fuzzy model that serves as the copilot for an RL model. Herein, the fuzzy model is logically configured based on human expert knowledge. As a result, this study combines human expertise with learning-driven approaches. This research also includes new sensitivity analyses to evaluate the performance of models in various conditions.

This study delivers significant implications for the industrial sector. With the ability to make decisions under uncertain conditions, the developed models have a great potential to effectively perform in industrial settings and deal with unpredictable factors that alter processes from ideal path. Through integrating sustainable factors in the HRC disassembly planning processes, this research allows manufacturers to customize processes based on their specific goals. In this manner, companies control the process workflows according to their individual requirements regarding the varying significance of cost, environmental, and social factors. In this case, the presented GUI offers a user-friendly tool built upon the models to manufacturers, enabling them to configure processes through the interactive dashboard. Furthermore, the proposed ergonomic risk assessment method is an effective and low-cost solution that can be easily implemented in industrial environments using cameras, eliminating the need for wearable sensors or other advanced equipments. Moreover, this study has substantial sustainability implications. In the environmental aspect, by reducing the energy consumed in a process, this model contributes to greenhouse gas reduction. Also, it effectively mitigates the use of materials and resource depletion by increasing the quality of recovered parts. From a social perspective, this model reduces the risk of injury to manual workers by considering ergonomic risk and human safety in the decision making process.

8.6 Limitations and future research

Difficulty and technical feasibility are inherently subjective concepts. In order to prove the concept, this research considers the technical feasibility values as an input, determined based on experts' knowledge. The difficulty values are also computed based on a weighted sum of the various characteristics of components, including weight and required force. Furthermore, we evaluated the models through the simulated scenarios instead of conducting experiments in industrial environments. Furthermore, human safety is a critical and sensitive concern in HRC disassembly processes. Although this research considers tasks involving toxic and hazardous materials as a safety criterion, human safety encompasses broader aspects. For example, identifying and maintaining a safe distance between humans and robots during task allocation is another important factor, not addressed in this research due to a lack of practical implementation in industrial settings.

The next step in this research is to deploy the proposed models in an industrial environment and evaluate their performance. Furthermore, more research could be conducted using advanced sensor devices or questionnaires to measure the difficulty of disassembly tasks. In addition, future studies can explore further safety aspects, such as maintaining safe distances and addressing other types of risks. These distances can be measured using image processing methods or advanced sensors. In addition, future studies could conduct a sensitivity analysis on fuzzy parameters by evaluating a model in both normal and fuzzy environments under different scenarios, thereby highlighting the impacts of fuzzifying uncertain parameters. Moreover, extended Reality (XR) and digital twin technologies offer significant potential to improve HRC disassembly processes in various ways. They can improve human skills in interacting with cobots, validating processes, simulating different scenarios, diagnose problems, and predict failures. Furthermore, future research can monitor body movements and postures to assess fatigue, ergonomics, and safety using these technologies. In addition, future research could explore a broader range of sustainability KPIs as optimization objectives or integrate the planning model with formal LCA methodologies.

REFERENCES

[1] M.-L. Lee *et al.*, "Task allocation and planning for product disassembly with human–robot collaboration," *Robotics and Computer-Integrated Manufacturing*, vol. 76, p. 102306, 2022.

[2] Z. Zhou *et al.*, "Disassembly sequence planning: Recent developments and future trends," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 233, no. 5, pp. 1450–1471, 2019.

[3] G. Liu *et al.*, "A machine learning approach for detecting fatigue during repetitive physical tasks," *Personal and Ubiquitous Computing*, pp. 1–18, 2023.

[4] T. Go *et al.*, "Genetically optimised disassembly sequence for automotive component reuse," *Expert Systems with Applications*, vol. 39, no. 5, pp. 5409–5417, 2012.

[5] X. F. Zhang *et al.*, "Parallel disassembly sequence planning for complex products based on fuzzy-rough sets," *The International Journal of Advanced Manufacturing Technology*, vol. 72, pp. 231–239, 2014.

[6] H.-J. Li, J. Jiang, and Y.-F. Wang, "Disassembly sequence planning based on extended interference matrix and genetic algorithm," *Computer Engineering and Design*, vol. 34, no. 3, pp. 1064–1068, 2013.

[7] Z. Chunming, "Optimization for disassemble sequence planning of electromechanical products during recycling process based on genetic algorithms," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 11, no. 4, pp. 107–114, 2016.

[8] W. Hui, X. Dong, and D. Guanghong, "A genetic algorithm for product disassembly sequence planning," *Neurocomputing*, vol. 71, no. 13-15, pp. 2720–2726, 2008.

[9] V. K. Gonnuru *et al.*, "Disassembly planning and sequencing for end-of-life products with rfid enriched information," *Robotics and Computer-Integrated Manufacturing*, vol. 29, no. 3, pp. 112–118, 2013.

[10] E. Kongar and S. M. Gupta, "Genetic algorithm for disassembly process planning," in *Environmentally Conscious Manufacturing II*, vol. 4569. SPIE, 2002, pp. 54–62.

[11] C. Chung and Q. Peng, "Evolutionary sequence planning for selective disassembly in demanufacturing," *International Journal of Computer Integrated Manufacturing*, vol. 19, no. 03, pp. 278–286, 2006.

[12] J. R. Li, L. P. Khoo, and S. B. Tor, "An object-oriented intelligent disassembly sequence planner for maintenance," *Computers in Industry*, vol. 56, no. 7, pp. 699–718, 2005.

[13] Z. Lu and Y. Sun, "Disassembly sequence planning of civil aircraft products for maintainability design," *Acta Aeronaut Astronaut Sin*, vol. 31, no. 1, pp. 143–150, 2010.

[14] H. Wu and H. Zuo, "Selective-disassembly sequence planning based on improved genetic algorithm," *Acta Aeronaut Astronaut Sin*, vol. 30, no. 5, pp. 952–958, 2009.

[15] F. Giudice and G. Fargione, "Disassembly planning of mechanical systems for service and recovery: a genetic algorithms based approach," *Journal of Intelligent Manufacturing*, vol. 18, pp. 313–329, 2007.

[16] L. Zhong *et al.*, "Disassembly sequence planning for maintenance based on metaheuristic method," *Aircraft Engineering and Aerospace Technology*, 2011.

[17] W.-C. Yeh, "Optimization of the disassembly sequencing problem on the basis of self-adaptive simplified swarm optimization," *IEEE transactions on systems, man, and cybernetics-part A: systems and humans*, vol. 42, no. 1, pp. 250–261, 2011.

[18] Wei-Chang Yeh, "Simplified swarm optimization in disassembly sequencing problems with learning effects," *Computers & Operations Research*, vol. 39, no. 9, pp. 2168–2177, 2012.

[19] W.-C. Yeh, C.-M. Lin, and S.-C. Wei, "Disassembly sequencing problems with stochasticprocessing time using simplified swarm optimization," *International Journal of Innovation, Management and Technology*, vol. 3, no. 3, p. 226, 2012.

[20] G. Jin, W. Li, and K. Xia, "Disassembly matrix for liquid crystal displays televisions," *Procedia CIRP*, vol. 11, pp. 357–362, 2013.

[21] W. Li *et al.*, "Selective disassembly planning for waste electrical and electronic equipment with case studies on liquid crystaldisplays," *Robotics and Computer-Integrated Manufacturing*, vol. 29, no. 4, pp. 248–260, 2013.

[22] Weidong. Li *et al.*, "A distributed service of selective disassembly planning for waste electrical and electronic equipment with case studies on liquid crystal display," *Cloud*

Manufacturing: Distributed Computing Technologies for Global and Sustainable Manufacturing, pp. 23–47, 2013.

[23] J. Xu, S.-Y. Zhang, and S.-M. Fei, "Product remanufacture disassembly planning based on adaptive particle swarm optimization algorithm," *Journal of Zhejiang University (Engineering Science)*, vol. 45, no. 10, pp. 1746–1752, 2011.

[24] X. Junfang, "Planning of selective disassembly sequence based on ant colony optimization algorithm," *JOURNAL OF COMPUTER AIDED DESIGN AND COMPUTER GRAPHICS*, vol. 19, no. 6, p. 742, 2007.

[25] M. Kheder, M. Trigui, and N. Aifaoui, "Optimization of disassembly sequence planning for preventive maintenance," *The International Journal of Advanced Manufacturing Technology*, vol. 90, pp. 1337–1349, 2017.

[26] X. Mi *et al.*, "Research and implementation on visualization system of disassembly sequence planning based on ant colony algorithm," in *Proceedings of the 2011 15th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. IEEE, 2011, pp. 581–585.

[27] X. Liu *et al.*, "Disassembly sequence planning approach for product virtual maintenance based on improved max–min ant system," *The International Journal of Advanced Manufacturing Technology*, vol. 59, pp. 829–839, 2012.

[28] J. Wang *et al.*, "Intelligent selective disassembly using the ant colony algorithm," *Ai Edam*, vol. 17, no. 4, pp. 325–333, 2003.

[29] C. Lu *et al.*, "A multi-objective disassembly planning approach with ant colony optimization algorithm," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 222, no. 11, pp. 1465–1474, 2008.

[30] Y. Luo, Q. Peng, and P. Gu, "Integrated multi-layer representation and ant colony search for product selective disassembly planning," *Computers in Industry*, vol. 75, pp. 13–26, 2016.

[31] H. Shan *et al.*, "Ant colony optimization algorithm-based disassembly sequence planning," in *2007 International conference on mechatronics and automation*. IEEE, 2007, pp. 867–872.

[32] H. Wang *et al.*, "Ant colony optimization for disassembly sequence planning," in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 42584, 2006, pp. 635–641.

[33] J.-R. Li, S. Tor, and, and L. P. Khoo, "A hybrid disassembly sequence planning approach for maintenance," *J. Comput. Inf. Sci. Eng.*, vol. 2, no. 1, pp. 28–37, 2002.

[34] S. M. McGovern and S. M. Gupta, "A balancing method and genetic algorithm for disassembly line balancing," *European journal of operational research*, vol. 179, no. 3, pp. 692–708, 2007.

[35] J. Guo *et al.*, "A multi-population cooperative coevolution artificial bee colony algorithm for partial multi-robotic disassembly line balancing problem considering preventive maintenance scenarios," *Advanced Engineering Informatics*, vol. 62, p. 102750, 2024.

[36] Y. Fang *et al.*, "Multi-objective multi-fidelity optimisation for position-constrained human-robot collaborative disassembly planning," *International Journal of Production Research*, vol. 62, no. 11, pp. 3872–3889, 2024.

[37] Y. Hu *et al.*, "An ontology and rule-based method for human–robot collaborative disassembly planning in smart remanufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 89, p. 102766, 2024.

[38] S. Lou *et al.*, "A human-cyber-physical system enabled sequential disassembly planning approach for a human-robot collaboration cell in industry 5.0," *Robotics and Computer-Integrated Manufacturing*, vol. 87, p. 102706, 2024.

[39] X. Zhang *et al.*, "Multi-objective robust optimization over time for dynamic disassembly sequence planning," *International Journal of Precision Engineering and Manufacturing*, vol. 25, no. 1, pp. 111–130, 2024.

[40] D. Grochowski and Y. Tang, "A machine learning approach for optimal disassembly planning," *International Journal of Computer Integrated Manufacturing*, vol. 22, no. 4, pp. 374–383, 2009.

[41] M. Godichaud, F. Pérès, and A. Tchangani, "Disassembly process planning using bayesian network," in *Engineering Asset Lifecycle Management: Proceedings of the 4th World Congress on Engineering Asset Management (WCEAM 2009), 28-30 September 2009*. Springer, 2010, pp. 280–287.

[42] J. Xiao *et al.*, "Dynamic bayesian network-based disassembly sequencing optimization for electric vehicle battery," *CIRP Journal of Manufacturing Science and Technology*, vol. 38, pp. 824–835, 2022.

[43] H. M. Kulkarni, "Reinforcement learning approach for disassembly," 2010.

[44] S. A. Reveliotis, "Uncertainty management in optimal disassembly planning through learning-based strategies," *Iie Transactions*, vol. 39, no. 6, pp. 645–658, 2007.

[45] H. Mao, Z. Liu, and C. Qiu, "Adaptive disassembly sequence planning for vr maintenance training via deep reinforcement learning," *The International Journal of Advanced Manufacturing Technology*, pp. 1–10, 2021.

[46] Z. Chen *et al.*, "Disassembly sequence planning for target parts of end-of-life smartphones using q-learning algorithm," *Procedia CIRP*, vol. 116, pp. 684–689, 2023.

[47] Y. Ren *et al.*, "A self-adaptive learning approach for uncertain disassembly planning based on extended petri net," *IEEE Transactions on Industrial Informatics*, 2023.

[48] X. Zhao *et al.*, "Reinforcement learning-based selective disassembly sequence planning for the end-of-life products with structure uncertainty," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7807–7814, 2021.

[49] M. Neves and P. Neto, "Deep reinforcement learning applied to an assembly sequence planning problem with user preferences," *The International Journal of Advanced Manufacturing Technology*, vol. 122, no. 11, pp. 4235–4245, 2022.

[50] B. Calli *et al.*, "Benchmarking in manipulation research: Using the yale-cmu-berkeley object and model set," *IEEE Robotics & Automation Magazine*, vol. 22, no. 3, pp. 36–52, 2015.

[51] Calli *et al.*, "The ycb object and model set: Towards common benchmarks for manipulation research," in *2015 international conference on advanced robotics (ICAR)*. IEEE, 2015, pp. 510–517.

[52] R. S. Tabar *et al.*, "Selective disassembly planning considering process capability and component quality utilizing reinforcement learning," *Procedia CIRP*, vol. 121, pp. 1–6, 2024.

[53] M. Bdiwi, A. Rashid, and M. Putz, "Autonomous disassembly of electric vehicle motors based on robot cognition," in *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2016, pp. 2500–2505.

[54] C. Harris, M. Stephens *et al.*, "A combined corner and edge detector," in *Alvey vision conference*, vol. 15, no. 50. Manchester, UK, 1988, pp. 10–5244.

[55] D. P. Brogan, N. M. DiFilippo, and M. K. Jouaneh, "Deep learning computer vision for robotic disassembly and servicing applications," *Array*, vol. 12, p. 100094, 2021.

[56] X. Zhang *et al.*, "Unsupervised human activity recognition learning for disassembly tasks," *IEEE Transactions on Industrial Informatics*, 2023.

[57] Zhang *et al.*, "Automatic screw detection and tool recommendation system for robotic disassembly," *Journal of Manufacturing Science and Engineering*, vol. 145, no. 3, p. 031008, 2023.

[58] A. Rehman *et al.*, "Globalization and renewable energy use: how are they contributing to upsurge the co2 emissions? a global perspective," *Environmental Science and Pollution Research*, vol. 30, no. 4, pp. 9699–9712, 2023.

[59] K. Wang *et al.*, "Energy-efficient robotic parallel disassembly sequence planning for end-of-life products," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 2, pp. 1277–1285, 2021.

[60] N. Hartono, F. J. Ramírez, and D. Pham, "Optimisation of robotic disassembly plans using the bees algorithm," *Robotics and Computer-Integrated Manufacturing*, vol. 78, p. 102411, 2022.

[61] B. Zhou and Q. Wu, "Decomposition-based bi-objective optimization for sustainable robotic assembly line balancing problems," *Journal of Manufacturing Systems*, vol. 55, pp. 30–43, 2020.

[62] J. Chen and X. Jia, "Energy-efficient integration of assembly line balancing and part feeding with a modified genetic algorithm," *The International Journal of Advanced Manufacturing Technology*, vol. 121, no. 3-4, pp. 2257–2278, 2022.

[63] Y. Fang *et al.*, "Multi-objective evolutionary simulated annealing optimisation for mixed-model multi-robotic disassembly line balancing with interval processing time," *International Journal of Production Research*, vol. 58, no. 3, pp. 846–862, 2020.

[64] M. Chand and C. Ravi, "A state-of-the-art literature survey on artificial intelligence techniques for disassembly sequence planning," *CIRP Journal of Manufacturing Science and Technology*, vol. 41, pp. 292–310, 2023.

[65] D. Mukherjee *et al.*, "A survey of robot learning strategies for human-robot collaboration in industrial settings," *Robotics and Computer-Integrated Manufacturing*, vol. 73, p. 102231, 2022.

[66] T. Koch, "Approach for an automated safety configuration for robot applications," *Procedia CIRP*, vol. 84, pp. 896–901, 2019.

[67] P. Bobka *et al.*, "Simulation platform to investigate safe operation of human-robot collaboration systems," *Procedia CIRP*, vol. 44, pp. 187–192, 2016.

[68] M. Faccio, I. Granata, and R. Minto, "Task allocation model for human-robot collaboration with variable cobot speed," *Journal of Intelligent Manufacturing*, pp. 1–14, 2023.

[69] K. E. Stecke and M. Mokhtarzadeh, "Balancing collaborative human–robot assembly lines to optimise cycle time and ergonomic risk," *International Journal of Production Research*, vol. 60, no. 1, pp. 25–47, 2022.

[70] D. Battini *et al.*, "Preventing ergonomic risks with integrated planning on assembly line balancing and parts feeding," *International Journal of Production Research*, vol. 55, no. 24, pp. 7452–7472, 2017.

[71] R. Ozdemir *et al.*, "Fuzzy multi-objective model for assembly line balancing with ergonomic risks consideration," *International Journal of Production Economics*, vol. 239, p. 108188, 2021.

[72] Y. Chen *et al.*, "Human workload and ergonomics during human-robot collaborative electronic waste disassembly," in *2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS)*. IEEE, 2022, pp. 1–6.

[73] S. Lou *et al.*, "Personalized disassembly sequence planning for a human–robot hybrid disassembly cell," *IEEE Transactions on Industrial Informatics*, 2024.

[74] H. Alrufaifi, B. Kumar, and J. L. Rickli, "Automated contact and non-contact constraint generation for disassembly feasibility and planning," *Procedia CIRP*, vol. 80, pp. 548–553, 2019.

[75] J. L. Rickli and J. A. Camelio, "Multi-objective partial disassembly optimization based on sequence feasibility," *Journal of manufacturing systems*, vol. 32, no. 1, pp. 281–293, 2013.

[76] J. Rickli and J. Camelio, "Partial disassembly sequencing considering acquired end-of-life product age distributions," *International Journal of Production Research*, vol. 52, no. 24, pp. 7496–7512, 2014.

[77] L. Guo *et al.*, "A grey wolf optimization algorithm for solving partial destructive disassembly line balancing problem consider feasibility evaluation and noise pollution," *Advanced Engineering Informatics*, vol. 60, p. 102418, 2024.

[78] L. Guo and X. Zhang, "Multi-granularity feasibility evaluation method of the partial destructive disassembly for an end-of-life product," *The International Journal of Advanced Manufacturing Technology*, vol. 116, pp. 3751–3764, 2021.

[79] E. Kroll, B. Beardsley, and A. Parulian, "A methodology to evaluate ease of disassembly for product recycling," *IIE transactions*, vol. 28, no. 10, pp. 837–846, 1996.

[80] E. Kroll and B. S. Carver, "Disassembly analysis through time estimation and other metrics," *Robotics and Computer-Integrated Manufacturing*, vol. 15, no. 3, pp. 191–200, 1999.

[81] K. B. Zandin, *MOST work measurement systems*. CRC press, 2002.

[82] P. Vanegas *et al.*, "Ease of disassembly of products to support circular economy strategies," *Resources, Conservation and Recycling*, vol. 135, pp. 323–334, 2018.

[83] C. Favi *et al.*, "A design for disassembly tool oriented to mechatronic product demanufacturing and recycling," *Advanced Engineering Informatics*, vol. 39, pp. 62–79, 2019.

[84] "Growth in ai and robotics research accelerates," https://www.nature.com/articles/d41586-022-03210-9, 2022.

[85] D. Rodríguez-Guerra *et al.*, "Human-robot interaction review: Challenges and solutions for modern industrial environments," *IEEE Access*, vol. 9, pp. 108 557–108 578, 2021.

[86] S. Parsa and M. Saadat, "Human-robot collaboration disassembly planning for end-of-life product disassembly process," *Robotics and Computer-Integrated Manufacturing*, vol. 71, p. 102170, 2021.

[87] M. Wurster *et al.*, "Modelling and condition-based control of a flexible and hybrid disassembly system with manual and autonomous workstations using reinforcement learning," *Journal of Intelligent Manufacturing*, vol. 33, no. 2, pp. 575–591, 2022.

[88] I. Belhadj, M. Aicha, and N. Aifaoui, "Product disassembly planning and task allocation based on human and robot collaboration," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 16, no. 2, pp. 803–819, 2022.

[89] W. Qu *et al.*, "Adaptive planning of human–robot collaborative disassembly for end-of-life lithium-ion batteries based on digital twin," *Journal of Intelligent Manufacturing*, pp. 1–23, 2023.

[90] Q. Liu *et al.*, "Human-robot collaboration in disassembly for sustainable manufacturing," *International Journal of Production Research*, vol. 57, no. 12, pp. 4027–4044, 2019.

[91] G. Yuan *et al.*, "A new heuristic algorithm based on multi-criteria resilience assessment of human–robot collaboration disassembly for supporting spent lithium-ion battery recycling," *Engineering Applications of Artificial Intelligence*, vol. 126, p. 106878, 2023.

[92] R. Zhang *et al.*, "A reinforcement learning method for human-robot collaboration in assembly tasks," *Robotics and Computer-Integrated Manufacturing*, vol. 73, p. 102227, 2022.

[93] A. Oulmane, A. A. Lakis, and N. Mureithi, "Automatic fault diagnosis of rotating machinery," *European Journal of Mechanical Engineering Research*, vol. 3, no. 2, pp. 19–41, 2016.

[94] J. E. Sierra-Garcia and M. Santos, "Deep learning and fuzzy logic to implement a hybrid wind turbine pitch control," *Neural Computing and Applications*, pp. 1–15, 2021.

[95] C. El Hatri and J. Boumhidi, "Fuzzy deep learning based urban traffic incident detection," *Cognitive systems research*, vol. 50, pp. 206–213, 2018.

[96] B. Walek and P. Fajmon, "A hybrid recommender system for an online store using a fuzzy expert system," *Expert Systems with Applications*, vol. 212, p. 118565, 2023.

[97] Z. Xi and G. Panoutsos, "Interpretable machine learning: convolutional neural networks with rbf fuzzy logic classification rules," in *2018 International conference on intelligent systems (IS)*. IEEE, 2018, pp. 448–454.

[98] F. Fan and G. Wang, "Fuzzy logic interpretation of quadratic networks," *Neurocomputing*, vol. 374, pp. 10–21, 2020.

[99] K. Bölat and T. Kumbasar, "Interpreting variational autoencoders with fuzzy logic: A step towards interpretable deep learning based fuzzy classifiers," in *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2020, pp. 1–7.

[100] H. R. Berenji, "A reinforcement learning—based architecture for fuzzy logic control," *International Journal of Approximate Reasoning*, vol. 6, no. 2, pp. 267–292, 1992.

[101] P. Y. Glorennec and L. Jouffe, "Fuzzy q-learning," in *Proceedings of 6th international fuzzy systems conference*, vol. 2. IEEE, 1997, pp. 659–662.

[102] M. Goharimanesh, A. Mehrkish, and F. Janabi-Sharifi, "A fuzzy reinforcement learning approach for continuum robot control," *Journal of Intelligent & Robotic Systems*, vol. 100, pp. 809–826, 2020.

[103] L. Chen *et al.*, "Conditional dqn-based motion planning with fuzzy logic for autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 2966–2977, 2020.

[104] P. Kofinas, G. Vouros, and A. I. Dounis, "Energy management in solar microgrid via reinforcement learning using fuzzy reward," *Advances in building energy research*, vol. 12, no. 1, pp. 97–115, 2018.

[105] Y. Singh and N. Pal, "Reinforcement learning with fuzzified reward approach for mppt control of pv systems," *Sustainable Energy Technologies and Assessments*, vol. 48, p. 101665, 2021.

[106] M. Chen *et al.*, "Reinforcement learning-based control of nonlinear systems using lyapunov stability concept and fuzzy reward scheme," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 10, pp. 2059–2063, 2019.

[107] F. Alfaverh, M. Denai, and Y. Sun, "Demand response strategy based on reinforcement learning and fuzzy reasoning for home energy management," *IEEE access*, vol. 8, pp. 39 310–39 321, 2020.

[108] J. R. Duflou *et al.*, "Efficiency and feasibility of product disassembly: A case-based study," *CIRP annals*, vol. 57, no. 2, pp. 583–600, 2008.

[109] M. Kheder, M. Trigui, and N. Aifaoui, "Disassembly sequence planning based on a genetic algorithm," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 229, no. 12, pp. 2281–2290, 2015.

[110] C. Chung and Q. Peng, "A hybrid approach to selective-disassembly sequence planning for de-manufacturing and its implementation on the internet," *The International Journal of Advanced Manufacturing Technology*, vol. 30, no. 5, pp. 521–529, 2006.

[111] S. Yang and Z. Xu, "Intelligent scheduling and reconfiguration via deep reinforcement learning in smart manufacturing," *International Journal of Production Research*, vol. 60, no. 16, pp. 4936–4953, 2022.

[112] Y. Liu *et al.*, "Scheduling of decentralized robot services in cloud manufacturing with deep reinforcement learning," *Robotics and Computer-Integrated Manufacturing*, vol. 80, p. 102454, 2023.

[113] A. H. Sakr *et al.*, "Simulation and deep reinforcement learning for adaptive dispatching in semiconductor manufacturing systems," *Journal of Intelligent Manufacturing*, pp. 1–14, 2021.

[114] H. Wang *et al.*, "Dynamic inventory replenishment strategy for aerospace manufacturing supply chain: combining reinforcement learning and multi-agent simulation," *International Journal of Production Research*, vol. 60, no. 13, pp. 4117–4136, 2022.

[115] X. Xia *et al.*, "A multi-agent convolution deep reinforcement learning network for aero-engine fleet maintenance strategy optimization," *Journal of Manufacturing Systems*, vol. 68, pp. 410–425, 2023.

[116] T. Yu, J. Huang, and Q. Chang, "Optimizing task scheduling in human-robot collaboration with deep multi-agent reinforcement learning," *Journal of Manufacturing Systems*, vol. 60, pp. 487–499, 2021.

[117] C. Li *et al.*, "An ar-assisted deep reinforcement learning-based approach towards mutual-cognitive safe human-robot interaction," *Robotics and Computer-Integrated Manufacturing*, vol. 80, p. 102471, 2023.

[118] Z. Liu *et al.*, "Task-level decision-making for dynamic and stochastic human-robot collaboration based on dual agents deep reinforcement learning," *The International Journal of Advanced Manufacturing Technology*, vol. 115, no. 11-12, pp. 3533–3552, 2021.

[119] M. El-Shamouty *et al.*, "Towards safe human-robot collaboration using deep reinforcement learning," in *2020 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2020, pp. 4899–4905.

[120] M. Dalle Mura and G. Dini, "Designing assembly lines with humans and collaborative robots: A genetic approach," *CIRP Annals*, vol. 68, no. 1, pp. 1–4, 2019.

[121] M. Dalle Mura and G. Dini, "Job rotation and human–robot collaboration for enhancing ergonomics in assembly lines by a genetic algorithm," *The International Journal of Advanced Manufacturing Technology*, pp. 1–14, 2022.

[122] Ö. F. Yılmaz *et al.*, "Tactical level strategies for multi-objective disassembly line balancing problem with multi-manned stations: an optimization model and solution approaches," *Annals of Operations Research*, vol. 319, no. 2, pp. 1793–1843, 2022.

[123] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[124] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[125] E. Álvarez-de-los Mozos, A. Rentería-Bilbao, and F. Díaz-Martín, "Weee recycling and circular economy assisted by collaborative robots," *Applied Sciences*, vol. 10, no. 14, p. 4800, 2020.

[126] S. Keivanpour, D. Ait-Kadi, and C. Mascle, "Automobile manufacturers' strategic choice in applying green practices: joint application of evolutionary game theory and fuzzy rule-based approach," *International Journal of Production Research*, vol. 55, no. 5, pp. 1312–1335, 2017.

[127] M. Chu and W. Chen, "Human-robot collaboration disassembly planning for end-of-life power batteries," *Journal of Manufacturing Systems*, vol. 69, pp. 271–291, 2023.

[128] L. Guo, Z. Zhang, and X. Zhang, "Human–robot collaborative partial destruction disassembly sequence planning method for end-of-life product driven by multi-failures," *Advanced Engineering Informatics*, vol. 55, p. 101821, 2023.

[129] S. Keivanpour, D. Ait Kadi, and C. Mascle, "End-of-life aircraft treatment in the context of sustainable development, lean management, and global business," *International Journal of Sustainable Transportation*, vol. 11, no. 5, pp. 357–380, 2017.

[130] S. Hezer and K. Yakup, "Balancing disassembly line under hazardous parts with precise and fuzzy goals," *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 11, no. 1, pp. 1–1, 2022.

[131] Z.-f. Zhang *et al.*, "A novel approach for parallel disassembly design based on a hybrid fuzzy-time model," *Journal of Zhejiang University-SCIENCE A*, vol. 16, no. 9, pp. 724–736, 2015.

[132] H.-P. Hsu, "A fuzzy knowledge-based disassembly process planning system based on fuzzy attributed and timed predicate/transition net," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 8, pp. 1800–1813, 2016.

[133] A. Allagui *et al.*, "Reinforcement learning for disassembly sequence planning optimization," *Computers in Industry*, vol. 151, p. 103992, 2023.

[134] W. Xu *et al.*, "Human-robot collaborative disassembly line balancing considering the safe strategy in remanufacturing," *Journal of Cleaner Production*, vol. 324, p. 129158, 2021.

[135] M. Chang, S. Ong, and A. Nee, "Approaches and challenges in product disassembly planning for sustainability," *Procedia Cirp*, vol. 60, pp. 506–511, 2017.

[136] A. Sartal *et al.*, "The sustainable manufacturing concept, evolution and opportunities within industry 4.0: A literature review," *Advances in Mechanical Engineering*, vol. 12, no. 5, p. 1687814020925232, 2020.

[137] Y. Kazancoglu and Y. D. Ozkan-Ozen, "Sustainable disassembly line balancing model based on triple bottom line," *International Journal of Production Research*, vol. 58, no. 14, pp. 4246–4266, 2020.

[138] B. Zhou and J. Bian, "Multi-mechanism-based modified bi-objective harris hawks optimization for sustainable robotic disassembly line balancing problems," *Engineering Applications of Artificial Intelligence*, vol. 116, p. 105479, 2022.

[139] C. Siew *et al.*, "Human-oriented maintenance and disassembly in sustainable manufacturing," *Computers & Industrial Engineering*, vol. 150, p. 106903, 2020.

[140] A. Amirnia and S. Keivanpour, "A context-aware real-time human-robot collaborating reinforcement learning-based disassembly planning model under uncertainty," *International Journal of Production Research*, pp. 1–22, 2023.

[141] A. Amirnia, E. Ghorbani, and S. Keivanpour, "Real-time video processing in fuzzy posture-based ergonomic analysis in a disassembly cell," in *International Conference on Intelligent and Fuzzy Systems*. Springer, 2024, pp. 247–256.

[142] X. Guo *et al.*, "Disassembly sequence planning: a survey," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 7, pp. 1308–1324, 2020.

[143] S. Keivanpour, "Toward joint application of fuzzy systems and augmented reality in aircraft disassembly," in *Intelligent and Fuzzy Techniques in Aviation 4.0: Theory and Applications*. Springer, 2021, pp. 265–280.

[144] L. McAtamney and E. N. Corlett, "Rula: a survey method for the investigation of work-related upper limb disorders," *Applied ergonomics*, vol. 24, no. 2, pp. 91–99, 1993.

[145] S. Hignett and L. McAtamney, "Rapid entire body assessment (reba)," *Applied ergonomics*, vol. 31, no. 2, pp. 201–205, 2000.

[146] M. Menanno *et al.*, "An ergonomic risk optimization system based on 3d human pose assessment and collaborative robot," 2023.

[147] L. Ma *et al.*, "A new muscle fatigue and recovery model and its ergonomics application in human simulation," *Virtual and Physical Prototyping*, vol. 5, no. 3, pp. 123–137, 2010.

[148] J. R. Potvin, "Predicting maximum acceptable efforts for repetitive tasks: an equation based on duty cycle," *Human factors*, vol. 54, no. 2, pp. 175–188, 2012.

[149] S. Gallagher and J. R. Heberger, "Examining the interaction of force and repetition on musculoskeletal disorder risk: a systematic literature review," *Human factors*, vol. 55, no. 1, pp. 108–124, 2013.

[150] Q. Lu *et al.*, "A hybrid metaheuristic algorithm for a profit-oriented and energy-efficient disassembly sequencing problem," *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101828, 2020.

[151] W. Xu *et al.*, "Disassembly sequence planning using discrete bees algorithm for human-robot collaboration in remanufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 62, p. 101860, 2020.

[152] J. Liu *et al.*, "Collaborative optimization of robotic disassembly sequence planning and robotic disassembly line balancing problem using improved discrete bees algorithm in remanufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101829, 2020.

[153] C. Zhan *et al.*, "Environment-oriented disassembly planning for end-of-life vehicle batteries based on an improved northern goshawk optimisation algorithm," *Environmental Science and Pollution Research*, vol. 30, no. 16, pp. 47 956–47 971, 2023.

[154] Y. Tao *et al.*, "Joint decision-making on automated disassembly system scheme selection and recovery route assignment using multi-objective meta-heuristic algorithm," *International Journal of Production Research*, vol. 57, no. 1, pp. 124–142, 2019.

APPENDIX A

ARTICLE 5: AI-DRIVEN EOL AIRCRAFT TREATMENT: A RESEARCH PERSPECTIVE

Authors: Ashkan Amirnia and Samira Keivanpour
Submitted on January 15, 2024, and published on July 31, 2024, in the *Intelligent Systems Conference (IntelliSys 2024)*, pages 371-391, DOI: 10.1007/978-3-031-66428-1_23.

1 Abstract

End-of-life (EoL) aircraft treatment is a crucial step in the circular manufacturing of the aircraft industry, delivering considerable economic and environmental benefits. Despite these advantages, complex and sensitive operations are a serious challenge in this field due to the high complexity of an aircraft structure and its constituent materials and composites. Poor operations can significantly reduce output quality and increase operational costs. Consequently, the value of the recovered materials may not exceed the operational costs. Artificial intelligence (AI) is a valuable tool for enhancing the effectiveness and sustainability of EoL aircraft management by automating and optimizing the identification, separation, processing, and transformation of materials. AI can also help to reduce waste and emissions, increase material recovery and reuse, and create new markets and jobs. This paper comprehensively reviews the AI applications in EoL aircraft treatment. It discusses the current state-of-the-art AI models across three domains: recycling, maintenance, and dismantling/disassembly. This article then carefully highlights the existing gaps based on the analyses. It also describes the possible directions for future research.

Keywords: Artificial intelligence; EoL aircraft; sustainable aircraft management; circular manufacturing

2 Introduction

The aerospace industry faces significant environmental and economic challenges due to the increasing amount of EoL aircraft that generate waste and pollution. To address this problem, circularity key performance indicators (KPIs) can help measure and monitor the progress of circular economy principles in the industry, such as designing out waste, keeping products and materials in use, and regenerating natural systems. The operationalization of circular economy in the industry with complex product configuration and supply chain is more challenging. To the best of our knowledge, there is no a common definition of circular economy and the related KPI at the product level. For example, research on the circular economy in the aerospace and railway industries is still in the infancy stage. Stavileci and Andersson [1] performed an analysis of the business model for the circular economy. They analyzed the case study of GKN Aerospace Sweden for evaluating different business models for material circularity (metal alloys) and the perspective of manufacturing technologies. Material scarcity, long production lead time, and the product life cycle are the main drivers for the circular economy in this context. The results of the study, based on the secondary data and interviews, reveal the need for more control over the product life cycle (in the design, manufacturing, and maintenance phases) and for using the recycled metal of the manufacturing process in the maintenance and production phases [2]. SNC-Lavalin Group also published a report on the circular economy overview in civil aerospace. They discussed that despite the small share of emissions from raw material extraction and aircraft production in comparison to the operation phase, for achieving the Net Zero goal, the circularity of material and energy should be

addressed during the whole life cycle of the products [3].

The synthesis of the literature highlights the following gaps: (1) the circularity of complex products requires deep analysis considering product configuration, life cycle assessment, and industrial context; (2) the new technologies in vehicles and lightweight constructions need novel approaches and strategies for the circularity of materials and energy, (3) in addition to circular economy analysis at the macro level, detailed case studies at the product level are required to provide an integrated approach to circular manufacturing practices. A systematic research approach for analyzing the circular economy in different industrial sectors and cross-industrial comparisons are required for identifying the best practices with a focus on material flows and processes. The decision context, applicable tools, and contribution of the advanced manufacturing technologies for boosting the circular economy should be studied. Current circular economy practices in different industrial sectors including Motor vehicles and parts, Telecoms, Industrial, Aerospace, and Railway should be analyzed at the company and product levels. Different aspects of the circular economy, including supply, product development, production, maintenance, and the EoL phase, will be taken into account. This will identify relevant metrics and highlight hotspots for integrating circularity in manufacturing.

AI refers to various methods that enable machines to reason and make decisions. These methods have great potential to solve complex and difficult problems. Nowadays, AI models, in particular machine learning (ML) algorithms, have wide applications in various fields, such as engineering, medicine, neuroscience, and business. ML algorithms can effectively extract non-linear relationships between data without requiring explicit rules, while they efficiently make decisions by analyzing historical data. Object recognition, abnormal detection, text analysis, and predicting patterns of moving objects are some of the classical applications of these algorithms. AI can play a crucial role in enabling circularity data and processing in the aircraft industry by enhancing the design of circular products and materials, optimizing complex recycling/reusing processes, and streamlining the infrastructure for product and material recovery.

This paper presents an in-depth review of AI applications in EoL aircraft treatment. It systematically examines previous studies in the field of AI in aircraft disassembly/dismantling, recycling, and maintenance to identify their strengths and weaknesses. This paper reveals the potential gaps in the literature and provides a comprehensive perspective for future research.

The rest of this paper is organized as follows. The following section discusses the background and literature review. Thereafter, Section 4 presents perspectives of AI applications in aircraft treatment. We discuss the potential gaps and the future research trends in Section 5. Finally, Section 6 concludes the article.

3 Background and literature review

This section reviews the background and relevant literature. First, it explains the context of EoL aircraft treatment, highlighting its benefits and challenges, such as the difficulty and complexity of operations involved. Subsequently, it explores AI applications in optimizing and enhancing the quality of complex and sensitive processes, such as cobotic disassembly, material composition detection, high precision cutting, and smart sorting. These processes not only hold potential for use in EoL aircraft treatment but also show the significant capability of AI for optimal decision-making within complex and challenging scenarios.

3.1 EoL aircraft treatment

The number of EoL aircraft is expected to increase significantly in the next decades due to the aging of the global fleet and the introduction of more environmentally friendly models. According to one estimate, about 12,000 aircraft will be retired by 2030 [4, 5]. The management of retired aircraft is essential considering the sustainability triple button lines. It can reduce the environmental impact of aviation by recovering valuable materials and components, minimizing waste generation, and saving energy and emissions. It creates economic opportunities for the aerospace industry by generating revenue from the sale of recovered parts and materials,

creating jobs in the recycling sector, and reducing the dependency on raw material imports. Moreover, it can enhance the social responsibility and reputation of aircraft manufacturers and operators by complying with environmental regulations and meeting the expectations of customers and stakeholders [6, 7]. According to Keivanpour et al. [8], dealing with EoL aircraft in a systematic way can benefit the aviation industry. The authors also proposed a holistic approach to EoL aircraft treatment that considers lean management, sustainable development, and global business [9].

However, there are many challenges associated with EoL aircraft recovery and closing the loop. Technical challenges are related to the lack of standardized and efficient methods and technologies for aircraft dismantling, sorting, testing, and certification. In addition, high costs and uncertainties of transporting, storing, and processing retired aircraft, and the competition from low-cost new parts and materials should be considered. The absence of specific and consistent regulations and standards for aircraft decommissioning, disassembly, and recycling, and the legal liabilities and risks involved are another limitation. The lack of coordination and collaboration among different stakeholders in the EoL aircraft supply chain, as well as the resistance to change are also the organizational challenges in this context and finally, the safe and sustainable disposal of retired aircraft, which contain hazardous materials and valuable resources, and the need to minimize the environmental impact and maximize the value of recovered parts and materials should be considered [4–10]. Furthermore, EoL aircraft treatment requires innovative solutions and multidisciplinary frameworks for addressing three pillars of sustainability. The projects related to EoL aircraft treatment are transdisciplinary because they involve different disciplines and players that need to collaborate and integrate their knowledge and perspectives. The projects need to address the technical, operational, and strategic aspects of EoL aircraft treatment, such as the business model, the market and industry, the knowledge management, and the performance measurement [4]. Therefore, it needs to develop advanced decision tools that can integrate different methods, and stakeholders in a coherent and flexible framework.

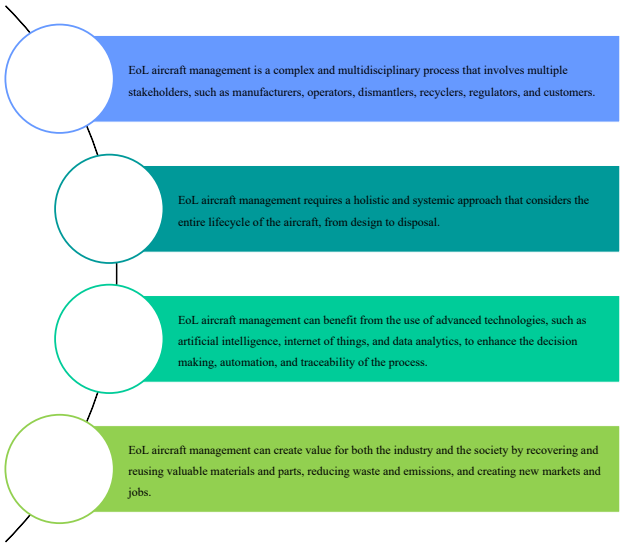


Figure 1: EoL aircraft context

Sustainability analysis is important in this context because it can help evaluate the impacts and benefits of EoL aircraft treatment from different perspectives. It can help identify best practices and strategies for EoL aircraft treatment that reduce the environmental footprint, increase social value, and optimize economic performance. It also provides monitoring and improving the performance of EoL aircraft treatment over time and across different locations. Sustainability analysis can also provide useful insights and feedback for decision-making and planning at both the strategic and managerial levels and help communicate and demonstrate the value proposition of EoL aircraft treatment to different stakeholders. The facilities for treatment of the aircraft at EoL in urban areas face some social challenges and opportunities. Keivanpour

et al. [2, 3, 8, 9, 11, 12] explored various aspects of EoL aircraft management, such as the impacts on design, modeling and optimization, applications of internet of things, and value chain analysis. The authors aimed to address the challenges and opportunities of sustainable development in the context of increasing aircraft retirements and recycling demands. They used different methods to analyze the technical, economic, environmental, and social factors that influence the EoL aircraft treatment and provided recommendations for improving the performance and efficiency of the EoL aircraft process. The highlighted points are summarized in Figure 1.

3.2 AI applications in high-value and complex products

In recent years, the use of AI algorithms—particularly ML and computer vision—in highly valuable, sensitive, and complex industrial processes has significantly increased. These algorithms efficiently optimize the processes, leading to increased quality and profit as well as reducing the related cost. As summarized in Figure 2, this section discusses AI applications in cobotic disassembly, material composite detection, sorting, and high-precision cutting, which are the most important sensitive and complex industrial processes.

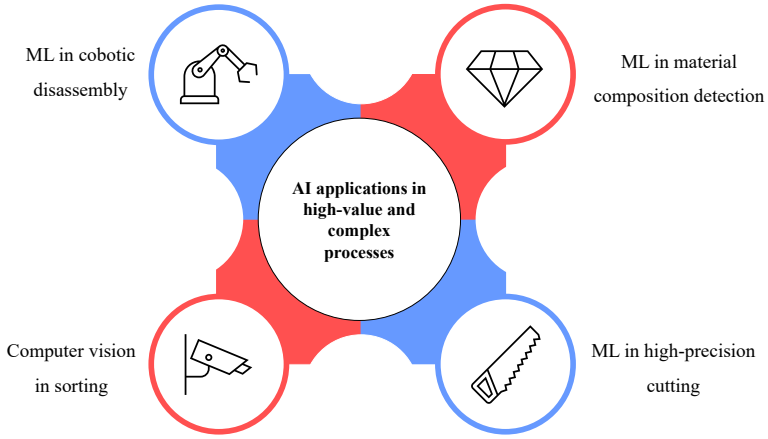


Figure 2: AI applications in highly complex, sensitive, and valuable industrial processes

3.2.1 ML in cobotic disassembly

In recent years, learning approaches have been widely used in cobotic disassembly, where a collaborative robot (cobot) works beside a human in a disassembly cell, shown in Figure 3. In this case, the purpose of using learning algorithms is to optimally plan sequential disassembly tasks with respect to various cost and profit factors. Learning models can handle uncertainties and after testing on several product configurations, the optimum disassembly sequence plan could be generated [13]. Goli et al. conducted a literature review on learning-based algorithms in disassembly planning [14]. Considering the different uncertainties of the returned products and components, the scheduled automated disassembly is challenging. Human-robot collaboration with flexibility can provide a solution to these difficulties. Various approaches have been put forward to generate an optimum task sequence in cobotic disassembly processes. Grochowski and Tang used an expert system based on Petri Net and a hybrid Bayesian network for disassembly planning [15]. Zude et al. studied recent developments in disassembly planning. According to the authors, maintenance conditions, MTTR, and tooling should be also considered in sequence planning [16]. In study [17], Parsa and Saadat introduced a new task classification and allocation in human-robot collaboration disassembly.

In order to obtain the physical precedence of tasks, the authors first model an EoL product architecture by an AND/OR graph approach. All disassembly operations are classified based on eight features, including the type of required tools, component size, weight, shape, accessibility, operation complexity, positioning,

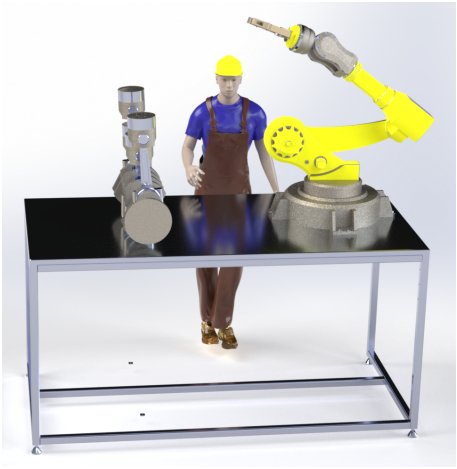


Figure 3: A human-robot collaboration disassembly cell

and operation force. In this case, these features represent the difficulty and complexity of each operation. The proposed planning model runs based on the Genetic algorithm (GA) that aims to find a near-optimal task sequence, resulting in the allocation of more complex and challenging tasks to human operators in addition to simple and repetitive tasks to cobots. The utilized GA includes several objectives in its fitness function, such as operation time, human operator change, and the non-targeted components term, which incentivizes the model to select the more important parts. Belhadj et al. also studied task allocation and disassembly planning with a cobot [18]. They used a CAD model and then generated the feasible sequence and selected the sequence with minimum time. However, they did not consider the product features and human and operator skills and constraints in sequence planning. Lee et al. developed task allocation and planning in a disassembly cell with a cobot. They also integrated the safety of the human operator into task allocation decisions [19]. In study [20], the authors have introduced a recursive cobotic disassembly framework including five steps: multi-model perception, multi-target cognition, decision-making, control and execution, and knowledge formation and evolution.

In the first step, different data types are collected from various resources, including sensors and cameras. By using different computer vision techniques in the second step, high-level features, including body skeletons, hand gestures, and tools positions, are extracted from the data. During the decision-making step, the next disassembly task is selected with respect to the dynamic situation of the process, represented by the extracted high-level features. Subsequently, a human or cobotic operator executes the selected task. In the last step, the learned knowledge is shared with other cobots through a cloud-based approach.

Despite supervised and unsupervised approaches, reinforcement learning (RL) algorithms can make decisions based on trial-and-error learning and the dynamic environmental context. Recently, several researchers have addressed cobotic disassembly planning by developing RL-based models. Chu and Chen [21] have presented a hybrid PSO with Q-learning for cobotic disassembly line balancing with respect to human safety, disassembly resources, deformation level of products, and task difficulty. In this case, the Q-learning method selects the most suitable optimizer to boost the PSO's efficiency. Allagui et al. [22] have developed a planning model based on Q-learning for human-robot collaboration disassembly. As the preprocessing step, they extracted CAD information from a product 3D model to define all feasible sequences along the x, y, and z angles. They have proposed a Q-tabular approach with a reward function that comprises the feasibility of movement along directions and a fitness function that includes several indicators, such as operation time, process direction and tool changes, and volumes of parts. In study [23], the authors have proposed a real time disassembly planning model based on a multi-agent RL framework. The proposed model allocates tasks to humans and cobots in real time with respect to the online status of operators and the product, as well as several criteria, including operation time, circularity, human skill-levels, and parts features. The authors

also represent the architecture of a product by using a graph-based approach.

3.2.2 ML in material composition detection

ML has been widely used in various fields, such as image recognition, natural language processing, and web search. In materials science, ML can predict material properties, optimize their design, and understand their behavior. One of the applications of ML is material composition detection. This aids in quality control, material characterization, and material recycling. Material composition detection can be done by using different experimental techniques, such as X-ray diffraction, X-ray fluorescence (XRF), Raman spectroscopy, or mass spectrometry. However, these techniques can be costly, time-consuming, or require expert knowledge to interpret the results. Gao et al. [24] developed an ML approach that combines data from high fidelity single-crystal measurements, low-fidelity polycrystalline measurements, and approximate simulations to predict the band gap of materials.

Another application of ML is the detection of the composition of complex materials, such as polymers or composites. Hence, ML can improve the accuracy and efficiency of material composition detection and facilitate the discovery and characterization of new materials. In the context of construction, Lu et al. have developed a computer vision model based on deep learning to recognize waste, such as rock, gravel, earth, and wood, by using a semantic segmentation approach [25]. In order to detect recycling status, Yang and Thung [26] have classified waste into six categories: paper, glass, plastic, metal, cardboard, and trash. In this research, the authors compared the performances of a CNN model and a model based on SVM with scale-invariant feature transform (SIFT). The experimental analysis illustrates that the SIFT+SVM model outperforms the CNN model. In another research study, [27] adopts CNNs to recognize Cu impurities in steel scrap. Here, the proposed approach achieves 90.6 % accuracy by applying Xception architecture. Recently, [28] uses magnetic induction spectroscopy (MIS) with ML for scrap metal classification. It concludes that SVM, random forest, and extra trees models deliver the best results for classifying non ferrous scrap metals. In study [29], the authors developed a waste-sorting and recycling classification model. To do that, they fused different features extracted by classical vision-based techniques. Following that, they applied an artificial neural network (ANN) for trash classification.

Moreover, in recent years, the use of ML methods to classify aluminum scraps has become more popular. Diaz-Romero et al. [30] have developed an approach based on deep learning and computer vision to classify cast and wrought aluminum scrap. In the first step, a region of interest (ROI) is selected by using basic image processing methods, such as morphological operations and Canny edge detection. Subsequently, a deep learning model classifies the objects located in ROI. Here, color and depth images are fused by using two different architectures. While the first architecture fuses color and depth images in the first layer, the second one simultaneously applies two separate sub-networks on both images and then, high-level extracted features are concatenated in the last layer. Using transfer learning techniques, the authors tested the proposed framework with five CNNs: AlexNet, DenseNet, RestNet18, SqueezeNet, and VGG16. Similarly, the study in [31] uses Laser-Induced Breakdown Spectroscopy (LIBS) in addition to color and depth images for the classification of aluminum scrap. The authors have developed two architectures based on deep learning. Each architecture includes three pipelines for processing the outputs of LIBS, RGB, and 3D cameras, and then, the results of these pipelines are fused in the last layer to form the final output.

3.2.3 ML in high-precision cutting

Cutting precision is important for various industrial applications, such as manufacturing, machining, fabrication, and dismantling. However, it can be affected by many factors, such as tool wear, material properties, environmental conditions, or human errors. AI can reduce human intervention and errors, increase productivity and quality, and save the time and cost of cutting. ML, computer vision, and image processing can help to improve cutting precision by monitoring, controlling, and optimizing the cutting process. Several

researchers have addressed cutting by using AI-based approaches. In study [32], the authors have developed a new image processing algorithm for detecting nodal points and faces in a honeycomb block. The proposed approach does not need high-resolution images and outperforms delivered algorithms by OpenCV in a real case study. Ma and Shao [33] have developed a real-time timber-cutting system based on deep learning and computer vision. The system initially detects defects in timber images captured from two vision sensors to find serviceable regions. After that, it classifies the timber as straight grain (SG), wavy grain (WG), wood stain (WS), or color difference (CD). By using a homography matrix, the system computes a cutting list containing cut points. Subsequently, the points are cut with respect to the list.

With the expansion of automated processes in the food industry, researchers have developed AI-based solutions for high-precision cutting in this field. Azarmdel et al. [34] have developed a model for detecting cutting points in trout fish. Without relying on learning-based approaches, the authors have used several classical image processing techniques, including thresholding, central and orientation detection, as well as morphological operations. Similarly, [35] proposes a sugarcane seed-cutting system based on basic conventional image processing techniques to recognize parts of sugarcane stalks. In another research, Mu et al. [36] have developed a full robotic system based on computer vision for cutting half-sheep. The authors initially extracted RGB and depth images of half-sheep by using an Azure Kinect sensor. Following that, they applied Deeplab v3+ model to detect the spine and ribs. Then, the cut curves are computed based on the detected spine and ribs coordinates. Finally, the proposed approach locates the 3D coordinates of the curves by exploiting the captured depth image. Next, Wu et al. have proposed a vision-based robotic system for cutting rachis in banana trees [37]. This system uses an improved version of YOLOv5 as well as classical image processing methods for identifying cut-off points. The system also obtains 3D information on these points using stereo vision. In the context of safety, a vision-based approach is recently introduced by Chiang et al. [38]. By reading images from a camera installed in the workplace, the model detects the hands of an operator working with a cutting machine. The machine turns off whenever the operator's hands are close to it. Table 1 shows AI technique applications in cutting precision.

Table 1: AI technique applications in cutting precision

Authors	Cutting process applications	AI techniques
[32]	Optimization in honeycomb blocks cutting	A proposed image processing approach detects nodal points and faces of honeycomb blocks
[33]	Automatic timbers cutting	Deep learning models can classify timbers in addition to detecting imperfect areas in them
[34]	Identifying cutting points in trout fish	A cascade of classical image processing techniques to locate cutting points and length
[35]	Sugarcane seed-cutting system	Classical image processing identifies cutting points in sugarcane stalks
[36]	Parameter optimization in half-sheep cutting	A CNN model detects spin and ribs, following that, a geometric analysis is used to compute 3D cut curves according to spin and ribs points as well as a depth image
[37]	Cutting rachis in banana trees	A CNN model recognizes bananas, and then, geometric analysis along with conventional image processing methods compute 3D spatial positions of cut-off points
[38]	Safety assessment in a cutting process	CNN detects a human operator's hands to ensure their safe distance from the cutting machine

3.2.4 AI in sorting

Sorting is separating items into different categories or groups based on some criteria. Sorting can be done manually or automatically, depending on the type and amount of items to be sorted. Sorting is applied for

quality control, recycling, inventory management, or product delivery. AI can help to improve the sorting process by using ML and computer vision techniques to identify and classify items based on their features, such as shape, color, size, texture, or label. In the recycling process, by using optical sensors and cameras different types of materials, such as paper, plastic, metal, or glass could be detected and sorted. Calaiaro [39] addressed how computer vision, ML, and robots can identify and sort recyclables based on their shapes, colors, and labels. Hence, AI can optimize the sorting parameters and reduce waste generation and disposal. It can increase the efficiency and accuracy of recycling sorting and reduce human intervention. However, there are still some challenges, such as data quality and availability, system integration and maintenance, ethical and social issues, and environmental impacts. Gondal et al. [40] compared different ML and deep learning models for classifying garbage waste images into glass, paper, metal, and plastic. They show that deep learning models perform better than traditional ML models. Fang et al. conducted a review on the use of AI for various aspects of municipal solid waste management, such as waste generation forecasting, waste segregation, waste treatment, and waste disposal. They also identify the gaps and challenges in this field [41]. As well, [42] presents a vision-based model for waste recognition and sorting, in which a deep neural network architecture is used for waste classification.

In addition, there is a growing body of literature developing AI-driven approaches for automated sorting in the food industry. For instance, Jeong et al. [43] have developed a ginseng sorting system that uses an ML model to classify ginseng into a class from three classes according to their shape and estimated weights. The authors compared the performance of SVM, MLP, and Inceptionv3, demonstrating that SVM outperforms other models. Further, Chen et al. [44] have introduced a robotic citrus sorting system containing a CNN to detect and classify defective fruits. Subsequently, the system employs the SORT algorithm and the Kalman filter for fruit tracking and predicting their trajectories, respectively. This approach enables a robotic arm to efficiently pick out defective fruits from the conveyor belt. Another study in [45] proposes a vision-based tomato grading system. Using several classical image processing techniques, the system initially extracts color, shape, and texture features from raw data. It subsequently detects defected and healthy tomatoes using different classifiers, including random forest, artificial neural network, and SVM with different kernels, including linear, quadratic, cubic, and RBF. Recently, the study in [46] introduces a vision-based model for grading-sorting citrus fruit. It also provides a hardware implementation to validate the model in a real scenario. The model efficiently meets the requirements of real-time processing with the use of a custom light deep neural network.

4 Perspectives of applications of AI in EoL aircraft treatment

In the same way as many other fields, the use of AI techniques in aircraft treatment is significantly growing. These techniques have a great potential to effectively enhance different aircraft treatment processes. As Figure 4 illustrates, the most applications of AI in aircraft treatment are in disassembly, dismantling, and recycling processes.

4.1 Maintenance

Many researchers have developed various ML and computer vision models to address different facets of aircraft maintenance. For example, [47] develops a vision-based system to detect faults in aircraft landing gear using a convolutional autoencoder. The difference between an input and the reconstructed images is defined as a measure to detect anomalies. If the difference is higher than a threshold, the system categorizes the corresponding image as an abnormal sample. Due to the presence of various potential factors, such as the lack of sufficient fault data, the images containing faults differ slightly from the corresponding reconstructed images. Hence, the proposed system cannot effectively classify normal and abnormal images. Furthermore, in study [48], the authors have proposed a learning-based approach to detect and localize impact damage on

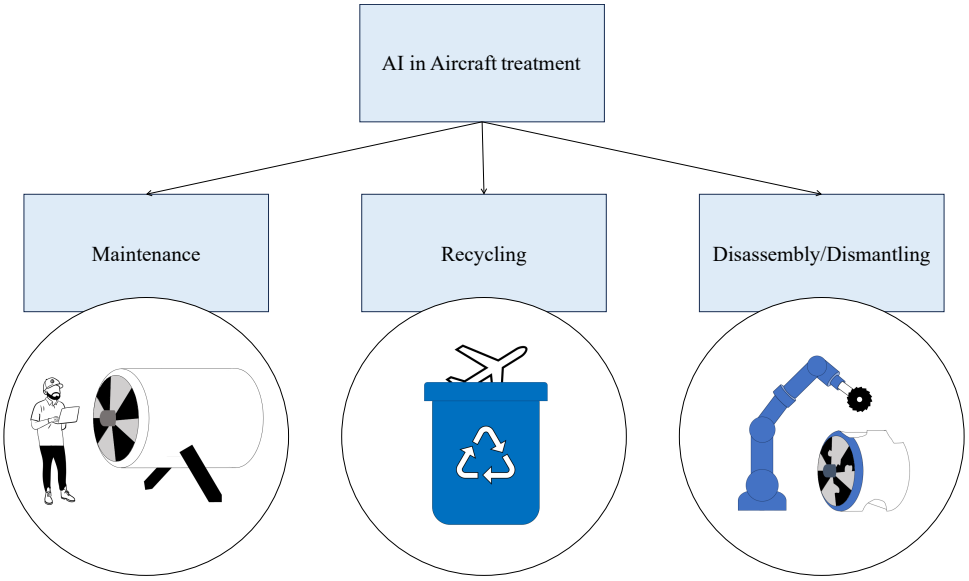


Figure 4: AI applications in aircraft treatment

aircraft composites based on acoustic emissions. To do that, they applied the random forest and a stacked autoencoder (SAE) for preprocessing acoustic emission data. In addition, the fast Fourier transform (FFT) was applied to acoustic emission data, and magnitudes of FFT were fed to SAE as an input. Since the acoustic emission waveform contains irrelevant information, using FFT magnitudes as the input plays the role of a representation learning mechanism and significantly improves SAE accuracy. The authors also investigated dimensionality reduction effects on the performance by removing unimportant features. The proposed approach was experimentally conducted on an aircraft elevator specimen in an offline setting. The analysis reveals that SAE delivers the best accuracy, while the random forest achieves accepted results with minimum computational time. In the same way, [49] presents a deep learning model for damage detection in aircraft fuselage. Herein, the problem of lacking data is effectively addressed by a transfer learning approach. Recently, Shafi et al. [50] have developed a real-time fault detection system based on deep learning for aircraft inspection. This system aims to reduce the cost of rework by identifying abnormal components in the early phase before assembling. Similarly, in [51], the authors have developed an intelligent system using Yolo v4 to recognize defects in the surface of landing gear components in real-time. They have trained the model based on their own collected images, delivering an experimental study to validate the proposed framework.

A number of researchers have predicted the remaining life of aircraft parts by ML algorithms. The research in [52] uses several linear and non-linear prediction models to estimate the remaining useful life of an aircraft engine. The authors used a dataset containing multivariate time-series observations representing the status of a simulated turbofan engine. To achieve the same goal, Kefalas et al. [53] have developed a learning-based method, including several steps, such as preprocessing, feature extraction, feature selection, and automatic modeling. Similarly, Asif et al. [54] have introduced an LSTM-based model to predict the remaining useful life of aircraft engines. They also proposed a new mathematical approach to identify the initial point of the engine failure and assign labels, resulting in the improved performance of the LSTM-based approach. Next, [55] applies a long-term differential method and the Fibonacci sliding window for preprocessing sensor data to compute long-term feature formation and short-term feature extension, respectively. Following that, a CatBoost model is used to predict the remaining life of aircraft engines based on these features. As well, Li et al. [56] used an ensemble learning approach containing several classifiers, such as autoregressive, relevance vector machine, and adaptive network-based fuzzy inference system, to predict the remaining life of an aircraft engine. Herein, the authors computed optimum weights for the classifiers by particle swarm optimization (PSO) and sequential quadratic programming (SQP) algorithms.

4.2 Recycling

AI-based approaches can optimize the processing and transformation methods based on material types, composition, and desired outputs. AI-based models can also track and verify the recycled materials and products throughout their life cycle. The tracking and traceability of materials used in aircraft, particularly with respect to recycling issues, can be investigated through blockchain as well as the joint application of AI and blockchain. A blockchain framework creates a history of material life cycles by tracking all activities relevant to them. AI algorithms have the potential to analyze and optimize corresponding processes based on data collected in the blockchain. This results in improving the recyclability of materials, and consequently reducing waste. Therefore, hybrid AI and blockchain systems can effectively enhance the productivity of material processes. A few researchers addressed the hybrid AI and blockchain applications with a particular focus on the aerospace industry. Abdulrahman et al. [57, 58] reviewed the applications of AI and blockchain in aerospace engineering. While the applications of AI and blockchain are becoming more and more popular in other fields, there has not yet been notable dedicated research for the aviation industry, promising a great opportunity for future studies.

Several attempts have been made to utilize AI in the context of space debris removal missions, which is related to EoL aircraft treatment. Yang et al. [59] have introduced an RL model for multi-debris active removal mission planning. They included the total time of a mission and the number of removed debris as the objectives in the reward function of the proposed model. Likewise, the study in [60] presents a decision-making model based on RL to plan a recycling rocket, aiming to remove space debris. The authors validated the presented framework by using an off-line analysis as well as comparing several deep RL models with different hyperparameters.

4.3 Disassembly/Dismantling

AI-driven models can optimize the disassembly/dismantling sequence and strategy based on the aircraft configuration, parts value, and customer requirements. These models also enable companies to monitor and control hazardous materials and fluids and prevent any leakage or contamination. A few studies have been published on developing AI-driven tools for aircraft disassembly. In addition, aircraft dismantling includes the process of disassembly, utilizing destructive or non-destructive approaches, or a combination of both, depending on the chosen EoL plan [61]. In order to facilitate the disassembly process of the EoL aircraft, Sabaghi et al. [62] have analyzed the impactful factors and their mutual interactions. Herein, a regression model based on a hybrid DOE-TOPSIS method computes disassemblability indices for the tasks. In this case, accessibility, quantity, and the diversities of the connections are the most influential factors in the tasks. Similarly, [61] develops a mathematical approach for aircraft disassembly concerning technical, economic, and environmental criteria. The proposed approach aims to analyze the ease of disassembly with respect to operation time, disassembly difficulty, and material compatibility for increasing recyclability. In this case, the difficulty score is obtained based on a linear summation of the required force for drilling, disengaging, and grinding. Furthermore, operation time consists of a standard operation time, tool calibration time, and an extra time for considering human fatigue. Material compatibility addresses a metric aiming to maximize output quality regarding material scarcity, alloying tolerance, and post-disassembly profitability. Recently, Blumel and Raatz [63] have introduced a learning-based approach based on analyzing the loosening torques and angle of rotation to detect and predict possible damage in aircraft disassembly. This strategy results in facilitating disassembly processes in addition to avoiding rework. The authors have implemented the introduced approach in a real-world scenario to validate its effectiveness through practical application. In another study, Keivanpour [64] presented a decision-making framework based on the integration of fuzzy logic and augmented reality for aircraft disassembly planning. Moreover, Yang et al. [7] reviewed the involvement of X-reality and lean in disassembly planning, with a particular focus on EoL aircraft components.

5 Discussions and future research direction

This section provides an overview of the methods used and discusses potential gaps within the existing literature. Figure 5 illustrates different applications of AI-driven approaches in aircraft EoL treatment. A review of previous methods shows the following findings.

Figure 5: Different applications of AI-driven approaches in aircraft EoL treatment

1. Numerous studies have been conducted in the field of computer vision applications for detecting damage and fault in airframes.
2. Many researchers have used sequential analysis approaches to predict the remaining life of aircraft parts.
3. Efforts have been undertaken to use ML models for planning space debris removal missions.
4. Relatively little research has been conducted to optimize disassembly/dismantling processes.
5. Limited research has explored material traceability through the combination of AI and blockchain.

Based on our findings, the following gaps exist in the existing literature and they can be addressed by future studies.

1. A few researchers have developed disassembly/dismantling models, dedicated to EoL aircraft. The feasibility of using AI to optimize the task allocation process in aircraft disassembly/dismantling using real or simulated data could be explored more by researchers.
2. Real-time analysis of the aircraft disassembly process by using computer vision-based methods can effectively provide valuable data for the decision-making process.
3. The development of a real-time decision-making model that can optimally identify opportunities for downcycling aircraft parts based on sustainability pillars and the quality of recovered parts.
4. The use of AI methods to optimize carbon fiber recycling process, playing an important role in EoL aircraft, has not been well-addressed in the literature. ML models can optimize the process based on recycling parameters, such as temperature and pressure. It is also possible to control the properties of recycled carbon fibers such as length, diameter and strength using ML.

5. The development of automatic waste management systems based on AI, dedicated for EoL aircraft, holds potential for the future exploration. Computer vision and ML models can identify and classify materials based on their features, such as shape, color, size, texture, or label. After recognizing and sorting, debris are placed in designated containers for recycling and reuse. Training computer vision models with the ability of recognizing materials and composites in sophisticated structure of aircraft is a potential gap in the literature.

6. The traceability of material used in aircraft with AI and blockchain is an emerging field and has not been adequately addressed in the literature.

6 Conclusion

The use of AI-based methods in various industries is increasingly expanding. AI models effectively help manufacturers optimize industrial processes, which results in increasing profit as well as reducing waste and consumed energy, leading to a movement towards sustainable manufacturing. The use of AI-based methods in EoL aircraft treatment is an emerging field in the aviation industry.

This article reviews the state-of-the-art research studies in EoL aircraft treatment. It thoroughly explains the serious problems and challenges within EoL aircraft in the aviation industry, as well as the importance of sustainable treatment and management for this case. The article also briefly discusses the economic and environmental benefits of EoL aircraft treatment. It comprehensively explains that AI-driven approaches have the great potential to enhance processes corresponding to EoL aircraft treatment. It then provides a systematic review of AI applications in challenging and high-value processes. Thereafter, it discusses AI driven methods in aircraft maintenance, recycling, and disassembly/dismantling as the most important processes related to EoL aircraft treatment.

The article aims to provide a comprehensive overview of AI applications, along with their challenges and opportunities, for EoL aircraft treatment. Overall, the use of AI-based methods in this field is fast-growing and leads to significant progress in optimizing relevant processes.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [grant number RGPIN-2020-05565].

References

[1] Shkrep Stavileci and David Andersson. An assessment of how circular economy can be implemented in the aerospace industry, 2015.

[2] Samira Keivanpour and D Ait Kadi. An integrated approach to analysis and modeling of end of life phase of the complex products. *IFAC-PapersOnLine*, 49(12):1892–1897, 2016.

[3] Samira Keivanpour, Christian Mascle, and Daoud Ait Kadi. A conceptual framework for value chain analysis of end of life aircraft treatment in the context of sustainable development. Technical report, SAE Technical Paper, 2014.

[4] Samira Keivanpour, Daoud Ait Kadi, and Christian Mascle. End-of-life aircraft treatment in the context of sustainable development, lean management, and global business. *International Journal of Sustainable Transportation*, 11(5):357–380, 2017.

[5] Samira Keivanpour, Daoud Ait Kadi, and Christian Mascle. End of life aircrafts recovery and green supply chain (a conceptual framework for addressing opportunities and challenges). *Management Research Review*, 38(10):1098–1124, 2015.

[6] Aircraft Fleet Recycling Association (AFRA). Afra website, 2025. Accessed: 30-Jan-2025.

[7] Yinong Yang, Samira Keivanpour, and Daniel Imbeau. Integrating x-reality and lean into end-of-life aircraft parts disassembly sequence planning: a critical review and research agenda. *The International Journal of Advanced Manufacturing Technology*, 127(5):2181–2210, 2023.

[8] Samira Keivanpour and Daoud Ait Kadi. Modelling end of life phase of the complex products: the case of end of life aircraft. *International Journal of Production Research*, 55(12):3577–3595, 2017.

[9] S Keivanpour and D Ait Kadi. The effect of "internet of things" on aircraft spare parts inventory management. *IFAC-PapersOnLine*, 52(13):2343–2347, 2019.

[10] Hamidreza Zahedi. *End-of-life efficient disassembly of complex structures using product and process focused approach*. Ecole Polytechnique, Montreal (Canada), 2016.

[11] Lhoussaine Ameknassi, Daoud Ait-Kadi, and Samira Keivanpour. Incorporating design for environment into product development process: an integrated approach. *IFAC-PapersOnLine*, 49(12):1460–1465, 2016.

[12] Salman Kimiagari, Samira Keivanpour, and Matti Haverila. Developing a high-performance clustering framework for global market segmentation and strategic profiling. *Journal of Strategic Marketing*, 29(2):93–116, 2021.

[13] Samira Keivanpour. *Circular Economy in Engineering Design and Production: Concepts, Methods, and Applications*. Springer Nature, 2023.

[14] Farzaneh Goli, Yongjing Wang, and Mozafar Saadat. Perspective of self-learning robotics for disassembly automation. In *2022 27th international conference on automation and computing (ICAC)*, pages 1–6. IEEE, 2022.

[15] DE Grochowski and Y Tang. A machine learning approach for optimal disassembly planning. *International Journal of Computer Integrated Manufacturing*, 22(4):374–383, 2009.

[16] Zude Zhou, Jiayi Liu, Duc Truong Pham, Wenjun Xu, F Javier Ramirez, Chunqian Ji, and Quan Liu. Disassembly sequence planning: Recent developments and future trends. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 233(5):1450–1471, 2019.

[17] Soran Parsa and Mozafar Saadat. Human-robot collaboration disassembly planning for end-of-life product disassembly process. *Robotics and Computer-Integrated Manufacturing*, 71:102170, 2021.

[18] Imen Belhadj, Mahdi Aicha, and Nizar Aifaoui. Product disassembly planning and task allocation based on human and robot collaboration. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 16(2):803–819, 2022.

[19] Meng-Lun Lee, Sara Behdad, Xiao Liang, and Minghui Zheng. Task allocation and planning for product disassembly with human–robot collaboration. *Robotics and Computer-Integrated Manufacturing*, 76:102306, 2022.

[20] Quan Liu, Zhihao Liu, Wenjun Xu, Quan Tang, Zude Zhou, and Duc Truong Pham. Human-robot collaboration in disassembly for sustainable manufacturing. *International Journal of Production Research*, 57(12):4027–4044, 2019.

[21] Mengling Chu and Weida Chen. Human-robot collaboration disassembly planning for end-of-life power batteries. *Journal of Manufacturing Systems*, 69:271–291, 2023.

[22] Amal Allagui, Imen Belhadj, Régis Plateaux, Moncef Hammadi, Olivia Penas, and Nizar Aifaoui. Reinforcement learning for disassembly sequence planning optimization. *Computers in Industry*, 151:103992, 2023.

[23] Ashkan Amirnia and Samira Keivanpour. A context-aware real-time human-robot collaborating reinforcement learning-based disassembly planning model under uncertainty. *International Journal of Production Research*, 62(11):3972–3993, 2024.

[24] Chan Gao, Xiaoyong Yang, Ming Jiang, Lixin Chen, Zhiwen Chen, and Chandra Veer Singh. Machine learning-enabled band gap prediction of monolayer transition metal chalcogenide alloys. *Physical Chemistry Chemical Physics*, 24(7):4653–4665, 2022.

[25] Weisheng Lu, Junjie Chen, and Fan Xue. Using computer vision to recognize composition of construction waste mixtures: A semantic segmentation approach. *Resources, Conservation and Recycling*, 178:106022, 2022.

[26] Mindy Yang and Gary Thung. Classification of trash for recyclability status. *CS229 project report*, 2016(1):3, 2016.

[27] Zhijiang Gao, S Sridhar, D Erik Spiller, and Patrick R Taylor. Applying improved optical recognition with machine learning on sorting cu impurities in steel scrap. *Journal of Sustainable Metallurgy*, 6:785–795, 2020.

[28] Kane C Williams, Michael D O'Toole, and Anthony J Peyton. Scrap metal classification using magnetic induction spectroscopy and machine vision. *IEEE Transactions on Instrumentation and Measurement*, 72:1–11, 2023.

[29] Mazin Abed Mohammed, Mahmood Jamal Abdulhasan, Nallapaneni Manoj Kumar, Karrar Hameed Abdulkareem, Salama A Mostafa, Mashael S Maashi, Layth Salman Khalid, Hayder Saadoon Abdulaali, and Shauhrat S Chopra. Automated waste-sorting and recycling classification using artificial neural network and features fusion: A digital-enabled circular economy vision for smart cities. *Multimedia tools and applications*, 82(25):39617–39632, 2023.

[30] Dillam Diaz-Romero, Wouter Sterkens, Simon Van den Eynde, Toon Goedeme, Wim Dewulf, and Jef Peeters. Deep learning computer vision for the separation of cast-and wrought-aluminum scrap. *Resources, Conservation and Recycling*, 172:105685, 2021.

[31] Dillam Díaz-Romero, Simon Van den Eynde, Isiah Zaplana, Chuangchuang Zhou, Wouter Sterkens, Toon Goedemé, and Jef Peeters. Classification of aluminum scrap by laser induced breakdown spectroscopy (libs) and rgb+ d image fusion using deep learning approaches. *Resources, Conservation and Recycling*, 190:106865, 2023.

[32] Maksim V Kubrikov, Mikhail V Saramud, and Margarita V Karaseva. Method for the optimal positioning of the cutter at the honeycomb block cutting applying computer vision. *IEEE Access*, 9:15548–15560, 2021.

[33] Hailong Ma and Mingwei Shao. Algorithm for automatic optimizing cross-cut saw based on computer vision techniques. *Engineering Research Express*, 5(4):045022, 2023.

[34] Hossein Azarmdel, Seyed Saeid Mohtasebi, Ali Jafari, and Alfredo Rosado Muñoz. Developing an orientation and cutting point determination algorithm for a trout fish processing system using machine vision. *Computers and Electronics in Agriculture*, 162:613–629, 2019.

[35] Deqiang Zhou, Yunlei Fan, Ganran Deng, Fengguang He, and Meili Wang. A new design of sugarcane seed cutting systems based on machine vision. *Computers and Electronics in Agriculture*, 175:105611, 2020.

[36] Shuai Mu, Haibo Qin, Jia Wei, Qingkang Wen, Sihan Liu, Shucai Wang, and Shengyong Xu. Robotic 3d vision-guided system for half-sheep cutting robot. *Mathematical Problems in Engineering*, 2020(1):1520686, 2020.

[37] Fengyun Wu, Jieli Duan, Puye Ai, Zhaoyi Chen, Zhou Yang, and Xiangjun Zou. Rachis detection and three-dimensional localization of cut off point for vision-based banana robot. *Computers and Electronics in Agriculture*, 198:107079, 2022.

[38] Chien-Lin Chiang, Ming-Yuan Peng, I-Long Lin, Yu-Wei Chou, Ji-Men Fong, and Yi-Yuan Chiang. Simple industrial cutting machine safety system based on computer vision. In *2023 IEEE 3rd International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB)*, pages 487–490. IEEE, 2023.

[39] Jason Calaiaro. Ai takes a dumpster dive: Computer-vision systems sort your recyclables at superhuman speed. *IEEE Spectrum*, 59(7):22–27, 2022.

[40] Ali Usman Gondal, Muhammad Imran Sadiq, Tariq Ali, Muhammad Irfan, Ahmad Shaf, Muhammad Aamir, Muhammad Shoaib, Adam Glowacz, Ryszard Tadeusiewicz, and Eliasz Kantoch. Real time multipurpose smart waste classification model for efficient recycling in smart cities using multilayer convolutional neural network and perceptron. *Sensors*, 21(14):4916, 2021.

[41] Bingbing Fang, Jiacheng Yu, Zhonghao Chen, Ahmed I Osman, Mohamed Farghali, Ikko Ihara, Essam H Hamza, David W Rooney, and Pow-Seng Yap. Artificial intelligence for waste management in smart cities: a review. *Environmental Chemistry Letters*, 21(4):1959–1989, 2023.

[42] Song Zhang, Yumiao Chen, Zhongliang Yang, and Hugh Gong. Computer vision based two-stage waste recognition-retrieval algorithm for waste classification. *Resources, Conservation and Recycling*, 169:105543, 2021.

[43] Seokhoon Jeong, Yong-Min Lee, and Sangjoon Lee. Development of an automatic sorting system for fresh ginsengs by image processing techniques. *Human-centric Computing and Information Sciences*, 7:1–13, 2017.

[44] Yaohui Chen, Xiaosong An, Shumin Gao, Shanjun Li, and Hanwen Kang. A deep learning-based vision system combining detection and tracking for fast on-line citrus sorting. *Frontiers in Plant Science*, 12:622062, 2021.

[45] David Ireri, Eisa Belal, Cedric Okinda, Nelson Makange, and Changying Ji. A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing. *Artificial Intelligence in Agriculture*, 2:28–37, 2019.

[46] Subir Kumar Chakraborty, A Subeesh, Kumkum Dubey, Dilip Jat, Narendra Singh Chandel, Rahul Potdar, NRNV Gowripathi Rao, and Deepak Kumar. Development of an optimally designed real-time automatic citrus fruit grading–sorting machine leveraging computer vision-based adaptive deep learning model. *Engineering Applications of Artificial Intelligence*, 120:105826, 2023.

[47] Falko Kähler, Ole Schmedemann, and Thorsten Schüppstuhl. Anomaly detection for industrial surface inspection: application in maintenance of aircraft components. *Procedia CIRP*, 107:246–251, 2022.

[48] Li Ai, Vafa Soltangharaei, Mahmoud Bayat, Michel Van Tooren, and Paul Ziehl. Detection of impact on aircraft composite structure using machine learning techniques. *Measurement Science and Technology*, 32(8):084013, 2021.

[49] Bruno Brandoli, André R de Geus, Jefferson R Souza, Gabriel Spadon, Amilcar Soares, Jose F Rodrigues-Jr, Jerzy Komorowski, and Stan Matwin. Aircraft fuselage corrosion detection using. 2022.

[50] Imran Shafi, Muhammad Fawad Mazhar, Anum Fatima, Roberto Marcelo Alvarez, Yini Miró, Julio César Martínez Espinosa, and Imran Ashraf. Deep learning-based real time defect detection for optimization of aircraft manufacturing and control performance. *Drones*, 7(1):31, 2023.

[51] Falko Kähler, Akshay CK Shetty, and Thorsten Schüppstuhl. Ai-based endpoint detection for surface defect removal on aircraft components. In *2023 IEEE/SICE International Symposium on System Integration (SII)*, pages 1–6. IEEE, 2023.

[52] Hussein A Taha, Ahmed H Sakr, and Soumaya Yacout. Aircraft engine remaining useful life prediction framework for industry 4.0. In *Proceedings of the 4th North America conference on Industrial Engineering and Operations Management, Toronto, ON, Canada*, pages 23–25, 2019.

[53] Marios Kefalas, Mitra Baratchi, Asteris Apostolidis, Dirk van den Herik, and Thomas Bäck. Automated machine learning for remaining useful life estimation of aircraft engines. In *2021 IEEE International Conference on Prognostics and Health Management (ICPHM)*, pages 1–9. IEEE, 2021.

[54] Owais Asif, Sajjad Ali Haider, Syed Rameez Naqvi, John FW Zaki, Kyung-Sup Kwak, and SM Riazul Islam. A deep learning model for remaining useful life prediction of aircraft turbofan engine on c-mapss dataset. *Ieee Access*, 10:95425–95440, 2022.

[55] Kunyuan Deng, Xiaoyong Zhang, Yijun Cheng, Zhiyong Zheng, Fu Jiang, Weirong Liu, and Jun Peng. A remaining useful life prediction method with long-short term feature processing for aircraft engines. *Applied Soft Computing*, 93:106344, 2020.

[56] Zhixiong Li, Kai Goebel, and Dazhong Wu. Degradation modeling and remaining useful life prediction of aircraft engines using ensemble learning. *Journal of Engineering for Gas Turbines and Power*, 141(4):041008, 2019.

[57] Yusra Abdulrahman, Edin Arnautović, Vladimir Parezanović, and Davor Svetinovic. Ai and blockchain synergy in aerospace engineering: an impact survey on operational efficiency and technological challenges. *IEEE Access*, 2023.

[58] Yusra Abdulrahman, Vladimir Parezanovic, and Davor Svetinovic. Ai-blockchain systems in aerospace engineering and management: Review and challenges. In *2022 30th Telecommunications Forum (TELFOR)*, pages 1–4. IEEE, 2022.

[59] Jianan Yang, Yu Hen Hu, Yong Liu, Xiaolei Hou, and Quan Pan. On the application of reinforcement learning in multi-debris active removal mission planning. In *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, pages 605–610. IEEE, 2019.

[60] Guangyin Jin, Jincai Huang, Yanghe Feng, Guangquan Cheng, Zhong Liu, and Qi Wang. Addressing the task of rocket recycling with deep reinforcement learning. In *Proceedings of the 6th International Conference on Information Technology: IoT and Smart City*, pages 284–290, 2018.

[61] Hamidreza Zahedi, Christian Mascle, and Pierre Baptiste. A multi-variable analysis of aircraft structure disassembly-a technico-economic approach to increase the recycling performance. *Sustainable Materials and Technologies*, 29:e00316, 2021.

[62] Mahdi Sabaghi, Christian Mascle, and Pierre Baptiste. Evaluation of products at design phase for an efficient disassembly at end-of-life. *Journal of Cleaner Production*, 116:177–186, 2016.

[63] Richard Blümel and Annika Raatz. Towards early damage detection during the disassembly of threaded fasteners using machine learning. *Procedia CIRP*, 116:480–485, 2023.

[64] Samira Keivanpour. Toward joint application of fuzzy systems and augmented reality in aircraft disassembly. In *Intelligent and Fuzzy Techniques in Aviation 4.0: Theory and Applications*, pages 265–280. Springer, 2021.