1-2021

Human-Robot Collaboration Using Commonsense Knowledge in Smart Manufacturing Contexts

Christopher Joseph Conti
Montclair State University

ABSTRACT

Human-robot collaboration (HRC), where humans and robots work together on specific tasks, is a growing part of smart manufacturing that entails artificial intelligence (AI) techniques in manufacturing processes. Robots need to be able to dynamically understand their working environments and human partners both accurately and quickly, as inaccurate or slow predictions can be dangerous to humans and collaborative tasks. To handle challenging environments, robots need to utilize commonsense knowledge (CSK), which is everyday knowledge about fundamental concepts, such as how basic objects interact with each other, what their properties are, and how they are associated. Human beings utilize CSK regularly, and robots can effectively collaborate with humans through it. This thesis outlines the fundamentals of CSK to provide prerequisite information and demonstrates how robots utilize it to collaborate with humans. The thesis also demonstrates the effectiveness of CSK and HRC through simulation studies and real-world human-robot collaboration experiments by deploying commonsense knowledge priorities and mathematical modeling for task optimization in robot action planning. Human-robot collaboration is compared with humans working without aid from robots. This thesis presents the results of this work along with a survey of relevant literature, as well as open issues for further research. To the best of our knowledge, ours is pioneering work on proposing a specific approach based on commonsense knowledge for human-robot collaboration in smart manufacturing.

Keywords: Artificial Intelligence, Big Data, Collaborative Robotics, Commonsense Knowledge, Human-Robot Interaction, Mathematical Modeling, Smart Manufacturing, Task Quality Optimization, Vehicle Assembly

MONTCLAIR STATE UNIVERSITY

Human-Robot Collaboration using Commonsense Knowledge in Smart Manufacturing Contexts

by

Christopher Joseph Conti

A Master's Thesis Submitted to the Faculty of

Montclair State University

In Partial Fulfilment of the Requirements

For the Degree of

Master of Science

January 2021

College of Science and Mathematics

Department of Computer Science

[redacted]

Dr. Aparna Varde

Thesis Sponsor

[redacted]

Dr. Weitian Wang

Committee Member

[redacted]

Dr. Vaibhav Anu

Committee Member

HUMAN-ROBOT COLLABORATION USING COMMONSENSE KNOWLEDGE IN SMART MANUFACTURING CONTEXTS

A Thesis

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

By

Christopher Joseph Conti

Montclair State University

Montclair, New Jersey

2021

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor Dr. Aparna Varde and co-advisor Dr. Weitian Wang for their guidance.

I would like to thank Dr. Vaibhav Anu for serving as my committee member.

I would like to thank the CS department for enrolling me as a Graduate Assistant.

I would like to thank the CS GPC Dr. Dawei Li and the CS Department Chair Dr. Constantine Coutras.

I would like to thank my family for their support.

TABLE OF CONTENTS

LIST OF FIGURES

THESIS TEXT

1. Chapter 1: Introduction

1.1. Background and Motivation

Human-robot collaboration, where humans and robots work together on tasks, is an important part of manufacturing. Collaborative robots have a variety of benefits over traditional robots, such as being capable of working alongside human beings in the same space and being designated to handle multiple tasks [1]–[5]. Adding additional space for robots or getting multiple types of robots for multiple tasks increases costs, which can make using collaborative robots less costly. Additionally, collaborative robots supplement humans rather than supplanting them [6]. Robots need to be able to collaborate with humans easily and while planning for dynamic real-world situations [7]. Commonsense knowledge (CSK), which is understanding objects, their properties, and how they relate to and interact with each other, is important for human robot collaboration. Humans have commonsense knowledge due to life experience, such as knowing that icy ground is slippery and should be walked on carefully. Robots have a difficult time with acquiring and using this knowledge, but need it in order to collaborate with humans [8].

1.2. Contributions of This Study

This thesis presents information on commonsense knowledge and human robot collaboration, specifically focusing on how the two areas are connected. This thesis also offers a simulated experiment demonstrating how commonsense knowledge can improve human-robot collaboration. This simulation involves a human and a robot collaborating to construct a vehicle from a given set of parts. Commonsense knowledge will be used to guide the robot's actions so constructing the vehicle is fast and simple for the human worker. In addition, this thesis displays how in person experiments confirmed the benefits of human robot collaboration for constructing a vehicle in the real world.

1.3. Thesis Structure

The thesis is structured as follows. First, prerequisite information about commonsense knowledge, human-robot collaboration and automated vehicles is presented in section 2. Afterwards, the proposed approach, the details of the methodology and the study's limitations are outlined in section 3. From there, the experiments and their results are presented in section 4. Lastly, conclusions and future work are presented in section 5. The rest of this article is organized into these sections.

2. Chapter 2: Related Work

2.1. Commonsense Knowledge

Commonsense knowledge focuses on real-world entities, their connections and how they interact [8]. Commonsense knowledge is obvious to humans and can be used even if it does not relate to the current task. For example, human drivers, especially in New Jersey, know that if one deer runs in front of a car, there may be others nearby and they should remain careful. This knowledge is not directly related to driving, but it can still aid with driving more effectively. Adding this type of knowledge to robots is difficult due to the extensive prior training required [9]. Robots are far more competent with factual knowledge, such as weather patterns, but they have more difficulty with forming connections, such as thunder preluding lightning and heavy rain [9]. Another common example is that robots can identify types of vehicles, but humans can tell based on the way a person drives if they need to be careful of that person. Adding commonsense knowledge to machines will help robots perform nearly as well as human beings.

Commonsense knowledge is used in various areas such as text understanding, computer vision and image processing, reasoning and planning [10]. Programs when comprehending texts typically handle individual words and short phrases, while CSK focuses on more of the entire passage, leading to greater accuracy [10]. Computer vision benefits from using commonsense properties such as continuously tracking objects even after they leave the frame in order to prevent surprises [10]. When robots work in real time environments, unpredicted events can occur. To handle these events, robots need to utilize commonsense reasoning, such as a catering robot not delivering a drink if the glass is empty. Robots are able to execute tasks more effectively with commonsense knowledge since they can handle corner cases more efficiently.

Commonsense knowledge bases (e.g. [8], [11], [12]) contain useful forms of knowledge. Semantics are defined as a specific field concepts, while pragmatics are defined as general concepts. An example of semantic knowledge is knowing the optimal foot position for a specific kick while pragmatic knowledge is knowing that continuously training outside in extreme heat is dangerous. Commonsense knowledge bases utilize both semantics and pragmatics. One knowledge base that uses CSK in the context of the Semantic Web is YAGO [13]. The Semantic Web attaches importance to words based on references to their context. YAGO is particularly effective not only due to having great breadth, it also utilizes multiple sources for knowledge, while many other systems only use one source [13]. YAGO gathers facts from sources such as WordNet, a knowledge base that connects related words, and Wikipedia, a well-known knowledge base that acts as an online encyclopedia [13]. Within YAGO, entities are connected to other entities through relationships and can be part of classes. Classes themselves can also act as entities, meaning classes can be related to each other and to other entities [13]. Relation instances can be connected to other relation instances to form connected relationships, such as the popularity of soccer players and the team they play for [13].

DBpedia is another substantial knowledge base, with 10 million entities and 1.46 billion facts [14]. The knowledge database collects and structures data from Wikipedia in order to form its structure [14]. The system parses data from Wikipedia, extracts useful information and outputs that information into a data storage system [14]. Users maintain the information while information from Wikipedia is mapped to the DBpedia ontology [14]. DBpedia is designed to update information based on pages on Wikipedia changing, allowing the stored knowledge to remain relevant and up-to-date with current events [14].

ConceptNet is another knowledge base that focuses primarily on connecting concepts and analyzing whole or major sections of texts while not focusing on determining the veracity of specific assertations [15]. ConceptNet's nodes are short fragments such as "moving forward" and "front entrance." Data analysis is made up of three phases. The first phase is the Extraction phase, where data is collected. The second phase is the Normalization phase, where data is normalized, i.e. words are made singular and determiners such as 'the' and 'an' are removed from syntactic constructs. The final phase is the Relaxation phase, where processing improves the network's connectivity and minimizes semantic gaps. ConceptNet can be applied in various fields, such as providing relevant translations, helping with understanding conversations and providing the meaning of words in a specific context [15].

WebChild, an initiative at the Max Planck Institute for Informatics, Germany, is another contemporary commonsense knowledge base [9]. WebChild stores and extracts commonsense concepts, properties and relationships from the Internet. The knowledge base has concepts described based on their, corresponding real-world domain, similar concepts, physical parts if applicable, its related activities, its relevant properties and its usual locations. Suitable pictures of commonsense concepts are provided as well. WebChild also allows users to view these concepts through the WebChild commonsense browser [9]. A partial snapshot of the interface is shown in Figure 1. The interface describes a concept by providing a picture of that concept, what domain it is part of, activities the concept is associated with and other attributes [9]. WebChild provides an overview for various commonsense concepts and can be applied to real world applications, such as object recognition systems, in order to improve their commonsense capabilities.

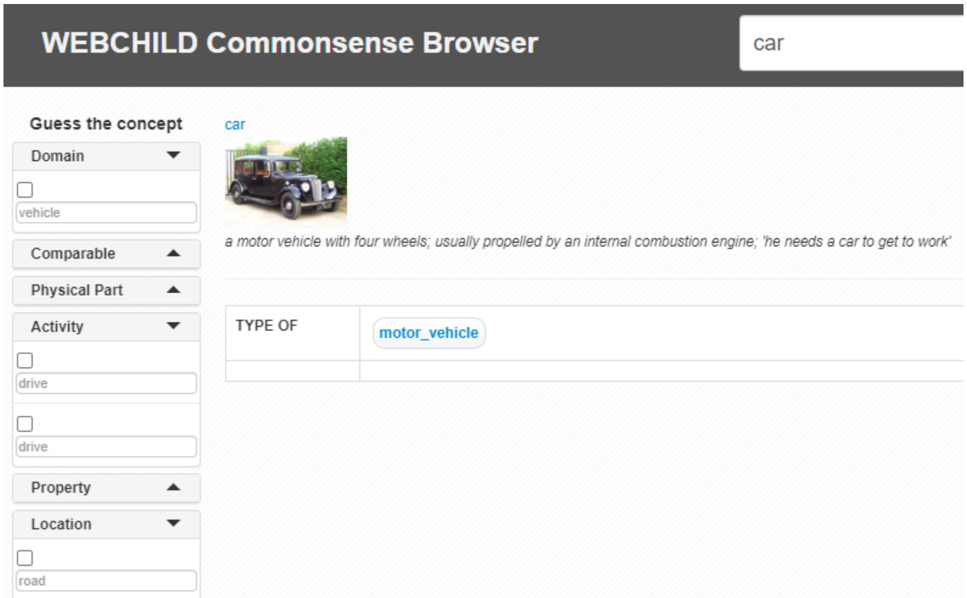

Figure 1: Partial snapshot of the WebChild commonsense browser

Translation tools are a frequently used knowledge base, demonstrated by the fact that Google Translate has over 500 million daily users, but are not without issues [16]. Translation tools often have trouble handling collocations, which are the correct form of colloquial expressions, such as ‘rat race’ instead of ‘rat rush’ [17]. Fixing unusual collocations can helps users search for information and with correcting errors in machine translated knowledge by using commonsense knowledge of collocations. CollOrder is a system that helps with that problem by detecting and correcting odd collocations [17]. CollOrder tags words by parts of speech and then searches its knowledge bases, which are made up of corrective native speaker English (e.g. the British National Corpus: BNC) collocations that correspond to the same parts of speech. CollOrder then ranks and filters suggestions and displays them based on their frequency. CollOrder can help users increase Google searches’ relevancy and help with developing writing aids for learns of ESL. Knowledge bases demonstrate how useful commonsense knowledge can be.

2.2. Human Robot Collaboration

Collaborative robots, which are designed for human robot collaboration, have several advantages over traditional robots. Traditional robots require extra equipment and guarding, resulting in greater costs and less flexibility [3]–[5]. Traditional robots can also only handle a pre-determined set of tasks, while collaborative robots can have the tasks they handle modified and expanded [1], [2]. In addition, collaborative robots support human strengths, such as judgement and adaptability, as they handle areas requiring strength and repetition [1].

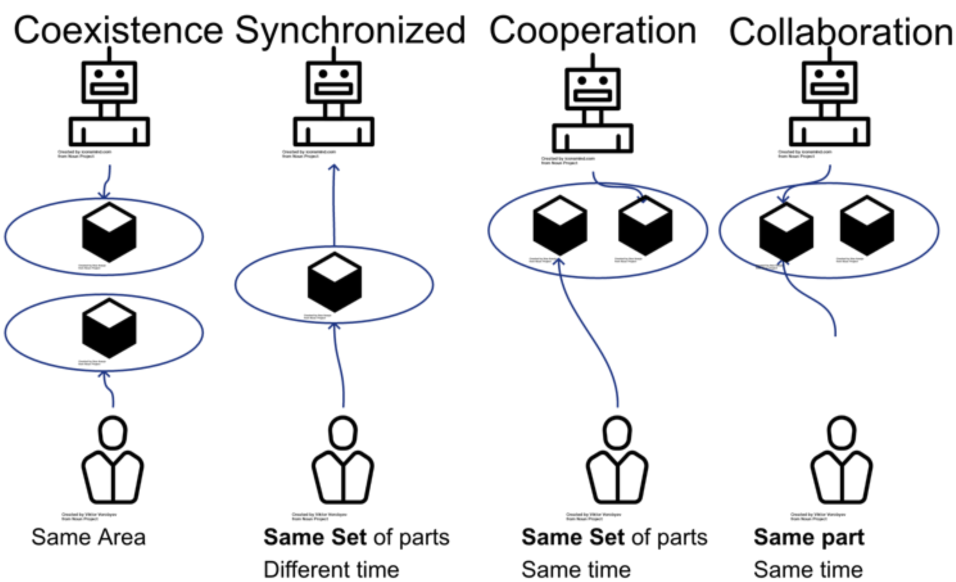

Figure 2: Types of human-robot collaboration
[Source: Icons for robot, human and part taken from "Noun project", see Appendix A][1].

The four main types of human-robot collaboration are coexistence, synchronized, cooperation and collaboration [1]. Cell collaboration involves robots and humans working in entirely separate spaces and not interacting. Coexistence entails humans and robots working in the same space without interacting. Synchronized collaboration entails humans and robots working in the same space, but at different times. Cooperation entails humans and robots working in the same space at the same time, but on different tasks. Finally, full collaboration entails humans and robots working on the same activities together. Full collaboration is not always the optimal form of collaboration; simpler forms of collaboration can be used in different situations. Human-robot collaboration is defined by various levels and can benefit companies.

Human-robot collaboration is useful since it allows robots to primarily handle certain tasks in order to increase safety and productivity [5]. Some of the main tasks that robots usually handle are welding, assembly and paint spraying. Robots handling these tasks benefits both human workers and companies utilizing robots. Robot welding results in better products while keeping humans safe and providing them with a better work life. Robot assembly decreases costs and increases consistency while moving humans away from monotonous work. Robot paint spraying results in greater consistency and productivity while protecting human workers from needing to repeatedly spray paint, which is dangerous [5]. Human-robot collaboration benefits companies with better results and cheaper costs and benefits human workers with more safety.

Robots need to be programmed to handle tasks, which involves programming them to react correctly to events in their environment and manage a sequence of waypoints [4]. However, this programming is time-consuming and often requires several iterations to create a sufficient system. Augmented reality makes testing a system easier through allowing an operator to see a simulation of the process and its results. The system would involve one operator handling a head mounted display, a camera, a hand-held input device and a wearable computer. The wearable computer handles various functions, including detecting key fixed points, rendering graphics and processing events. A tracking system follows the head mounted display and generates and places graphics in a 3D space based on operator actions. This allows the operator to simulate and view robot programming. Robot programming is around five times faster than traditional programming and operators find robot programming easier and more intuitive as well.

Programming commonsense knowledge in robots results into better performance since it connects spatial and temporal relationships between objects that make up activities [18]. Researchers at the University of Brennen, Germany are currently designing robot that utilizes commonsense knowledge for task execution. The system stores entities' 3D positions at a certain time in a four-dimensional vertex, providing the model with the positions of all objects at a certain time and a 3d image of the environment. The system also connects visual patterns detected from the image to spatial relations. For example, 2D and 3D shapes can be connected to objects’ size and distance. The system also allows for dividing tasks into simpler tasks that easier to handle. The commonsense knowledge features up this system make task execution easier for robots.

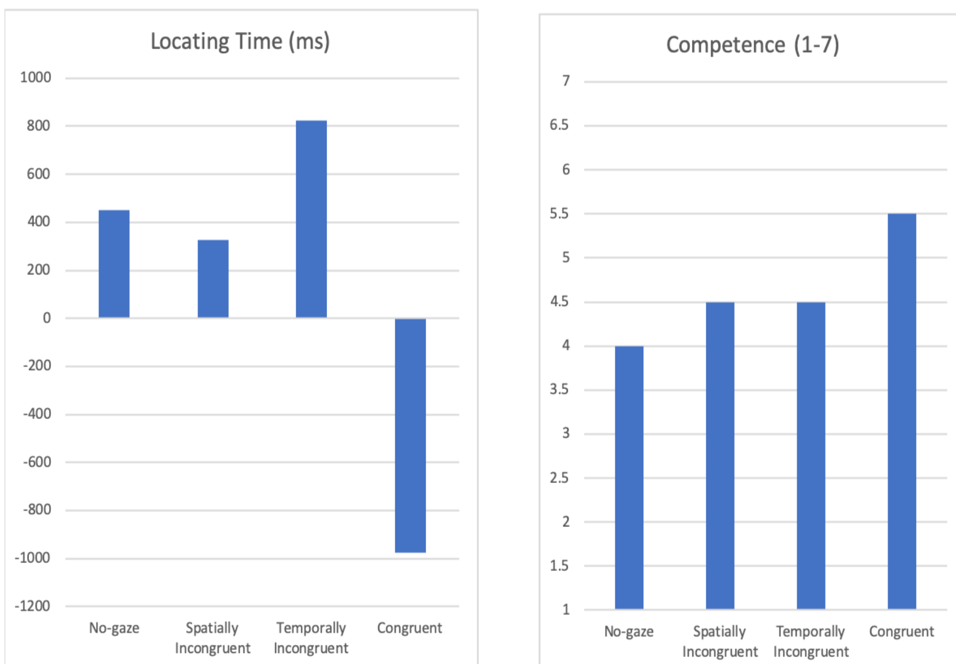

Figure 3: Testing results for different types of robot cues [19].

When humans communicate with each other while working on tasks, they send out and rely on cues. For robots to collaborate with humans, they need to be able read signals from humans while providing their own cues to humans to increase coordination [19]. One type of cue is joint attention, where two entities look at the same area. Researchers studied how different types of gaze cues from robots would affect joint attention. For this study, a robot asked participants to move objects with different colors to boxes of different colors. The robot provided three types of cues: congruent cues by referencing an object and then looking at it, temporally incongruent cues by looking at an object before referencing that object, and spatially incongruent cues by referencing an object and then looking at a different object. For control, the robot provided no cues. Results indicated that congruent gaze cues were the most useful since they vocally referenced objects followed by gazing at that object. Humans were able to most quickly locate objects and felt the most competent with vocal cues.

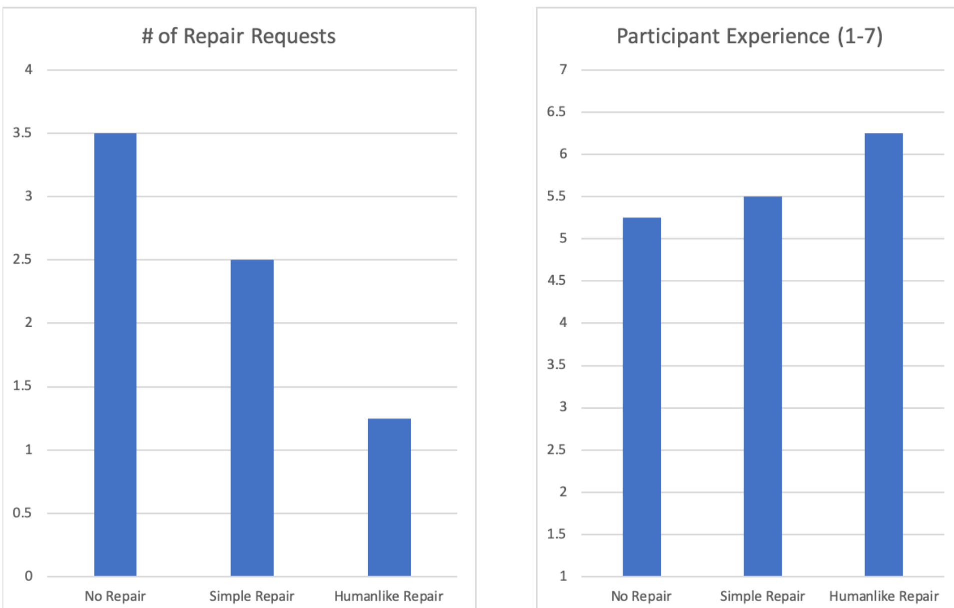

Figure 4: Testing results for different types of repair mechanisms [19].

Communication issues can occur when humans are cooperating, such as one person asking for more information, a person waiting and hesitating or a person incorrectly executing a task. When this occurs, people often repair issues by providing more information. Human robot collaboration can be improved by having robot provides these repairs when issues occur. A study tested how robots using different repair types would correspond to the total number of breakdowns in communication. Repairs can either be non-existent, where the robot only provides additional information once one task is completed, simple repair, where the robot only responds to yes or no questions and repeats instructions for other types of questions and humanlike repair, where the robot responds to questions in an appropriate, humanlike manner. Humanlike repair led to significantly fewer breakdowns and a better experience for participants. Adding humanlike communication to robots allows for better human-robot collaboration.

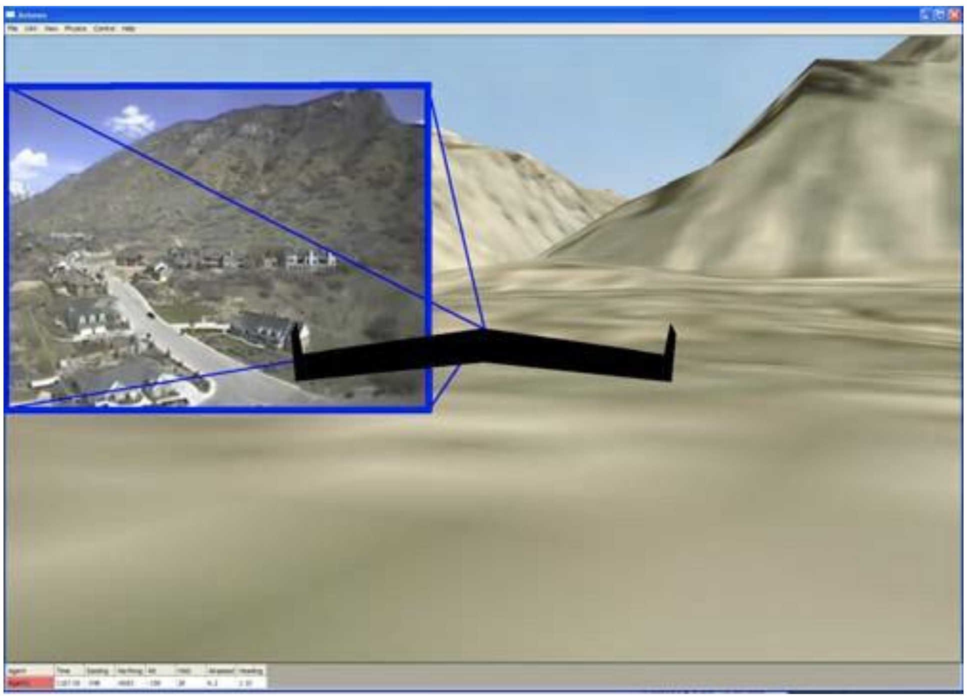

Figure 5: Augmented reality cockpit interface [20]

Human robot collaboration has grown more popular, with robots being used more frequently by non-specialist users [20]. Designing an effective interaction model is more important when working with non-specialists and video games can be used as a model. Video games provide information in a clear manner and are simple and enjoyable to control. An augmented reality cockpit interface demonstrates how a video game model works better than a virtual cockpit screen, which is the traditional model. The virtual cockpit screen provides several data points and a top-down view of navigation waypoints on a 2D screen. Rather than presenting the data in 2D, the augmented reality interfaced shows an outline of the aircraft along with live video from outside of the airplane. Whereas the virtual cockpit demands that several areas be focused on, the augmented reality interfaces only requires focusing on two areas. By presenting information in the simple and clean manner that videos game do, robots can more effectively support humans.

One of the parts of human robot collaboration that is most related to this thesis is that robots need to possess effective and customizable motion planning [21], [22]. Robots that can effectively plan are handle more tasks correctly with high productivity and low costs [23]. In order to aid planning, robots should have certain principles that determine their actions [24]. For example, robots should handle more dangerous objects so that humans are safer while assuring their actions follow the steps for the task. Robots can undertake actions that minimize a task cost determined from various real-world parameters while protecting human beings [24]. Implementing motion planning can improve human robot collaboration by aiding human workers while producing better results.

Although CSK-based interaction and effective communication are occurring, humans ultimately need to trust robots, i.e. have faith in them to collaborate with them, which makes the concept of *trust* very important [24]. If human users overly trust a robot, there may be issues when it is given too many tasks to handle. If users insufficiently trust a robot, its productivity and usefulness will be reduced. Human trust in robots is a feature to be managed such that the robot is highly effective.

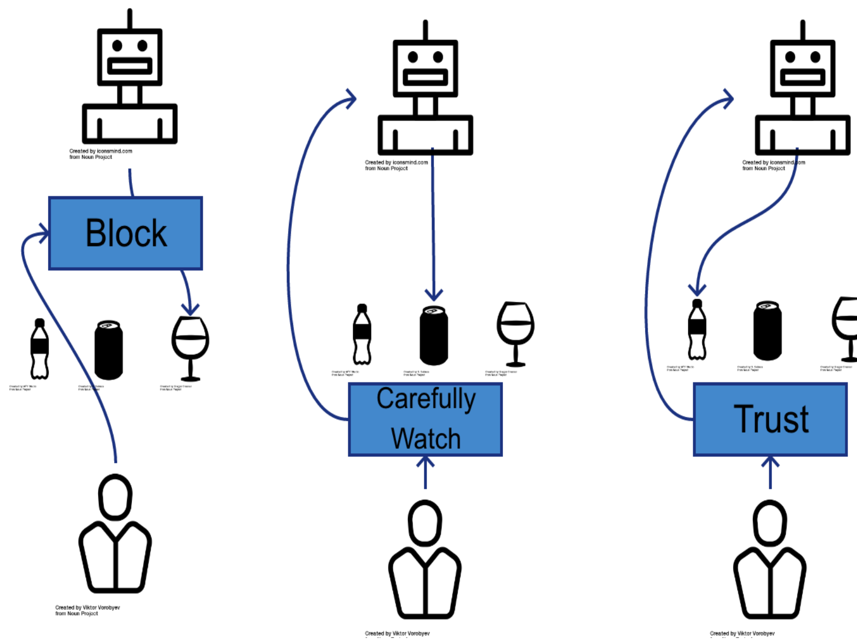

Figure 6: A robot removing objects from a table with a human collaborating: levels of trust (objects are bottle, can and glass, from left to right in each case) [Source: Icons of robot, human, bottle, can and glass taken from “Noun Project”, see References]

The robot can either ignore trust or focus on it. Interactions illustrated in Figure 6 display the levels of *estimated trust* between a human and a robot, adapted as found from a recent study. In this figure, the different iterations show the probable action the human will take based on the robot's action, with more trust being provided the less risky the action is. For example, in order to build trust, robots can start with low risk tasks so that they are trusted for higher risk tasks. In a study within this area, researchers had a person and a robot cooperate to clean a table, with bottles, cans and wine glasses placed on it [24]. Robots that focused primarily on maximizing the reward removed the wine glass first, causing more human intervention. Robots that focused on trust removed bottles before attempting to remove the wine glass, which caused the humans to be less likely to intervene. This would allow for removing the wine glass later when the human trusts the robot to do so. Detecting when a human is overly trusting a robot is important as well, and having the robot intentionally mismanage a task can be used to regain the focus of a human collaborator. By managing trust correctly, robots can better help humans with various tasks.

Robots use policies to determine what is to be done and where the actions are to be performed, based on the current world state. Robots are traditionally taught to handle tasks based on domain models and mathematical policies, but these approaches require defining the domain accurately. Domain experts are needed to develop domain-specific models. Rather than learning from experts or through precisely defined domains, robots can use *learning from demonstration (*LfD), where robots learn by watching another entity execute a task [25]. LfD is particularly beneficial since it does not require expert knowledge and ordinary people can demonstrate how to execute tasks. This makes LfD more flexible and intuitive than traditional robot teaching systems. There are several different LfD forms. The first is teleoperation, where the teacher operates the robot learner and the robot's sensors save the inputs. The second is shadowing, where the teacher executes a task and the robot attempts to mimic the teacher's actions while recording the task. The third is imitation, where the robot watches a non-identical entity perform a task while recording the task through its own sensors or sensors on an operator. The major difference between shadowing and imitation is that the robot is attempting to execute the task when shadowing and only watching the task for imitation. The different forms of learning from demonstration can be utilized to help robots learn how to execute the respective tasks. CSK can help improve robot learning since the robot will learn how to adapt to new situations and make decisions in a manner similar to human beings. It would bring such systems closer to the thresholds of human cognition.

Figure 7: Workers using a robot vehicle that acts as a platform for tree fruit farming [26].

Human robot collaboration has been utilized in agriculture, which is important since more food will need to be produced as the global population is expected to reach 9.1 billion by 2050, with 70 percent of that population living in urban areas [26]. Global food production will need to increase by 70%, which is where agricultural robots can aid human workers.

Robots are able to support human workers so they can be more efficient and this is demonstrated by a robot vehicle aiding with tree fruit farming [26]. The vehicles perform synchronized collaboration in Mule Mode and Pace Mode, where they assist with tasks while following workers, and perform tasks over a specified area, respectively. The vehicles in Scaffold Mode perform full collaboration, where the robot acts as a platform that humans can stand on for executing tasks. When compared to when human workers were using ladders to trim the trees, humans standing on the robots in Scaffold Mode were able to trim trees more than twice as quickly. Humans and robots can collaborate in order to increase food production, allowing for more people to have access to food.

2.3. Automated Vehicles

Smart mobility is an application of human-robot collaboration, where a robot completely or partially handles driving a vehicle. Automation is defined by several levels [27]. Level 0 of automation is no automation, where cruise control is greatest degree of automation. The first level is driver assistance, where dynamic cruise control and lane assistance are provided. The second level is partial automation, where assistance is provided for both controlling speed and steering. Up until the third level, the driver is primary entity handling the vehicle. The third level is conditional automation, where the robot can drive under ideal circumstances. The fourth level of automation is high automation, where vehicles can perfectly handle known use cases by themselves, but require a driver for unknown cases. The final level is full automation, where the vehicle is able to drive itself under any conditions. Autonomy is defined at this level, since it is where the vehicle makes decisions consistently.

Autonomous vehicles have uses cases outside of cars, as demonstrated by Unmanned Air Vehicles (UAV) developed by Brigham Young University [28]. Unmanned air vehicles are used due to their small size and inexpensive nature. However, they typically require a pilot. Brigham Young University’s UAVs are able to use autopilot to fly to a destination, while a human can provide instruction to the UAV if necessary. In order to reach their destination, the UAVs calculate a path made up of positions and times to reach each position. UAVs benefit from being able to fly continuously and not needing space for a person, further minimizing their size. They also do not need to be controlled by a handler due to their automated nature.

Vehicles require their AI to see what is occurring, process that information and act accordingly in order for full automation to occur [29]. Vehicles can use commonsense knowledge in order to handle environment processing and decision making, especially when handling unpredicted environments. Commonsense knowledge can aid with handling those issues [30], [31]. Since commonsense knowledge focuses on how objects are related to the context where they occur, it can be used to detect issues. For example, a vehicle's commonsense knowledge system knows that if another vehicle is quite close behind, the vehicle should move out of the way. Autonomous vehicles need utilize commonsense knowledge to analyze their environment in order to handle unpredictable situations safely.

While autonomous vehicles are beneficial, their usage removes the feeling of driving from people [32]. Autonomous vehicles can provide interaction through reading human hand gestures in order to improve the driving experience. Reading voice commands is a potential approach, but it is not as effective since often driving environments are loud and words need to be picked up clearly. Gestures such as turning, lane changing, increasing speed, decreasing speed, orienting the car and canceling inputs can allow for interaction between the driver and an autonomous vehicle. At the same time, these simple gestures provide an enjoyable driving experience while reducing the task workload. In addition, the driver feels they have some control over the vehicle. Automation has some issues that need to be considered when it is being implemented within vehicles. Adding gestures can provide a sense of control for humans while maintaining safety.

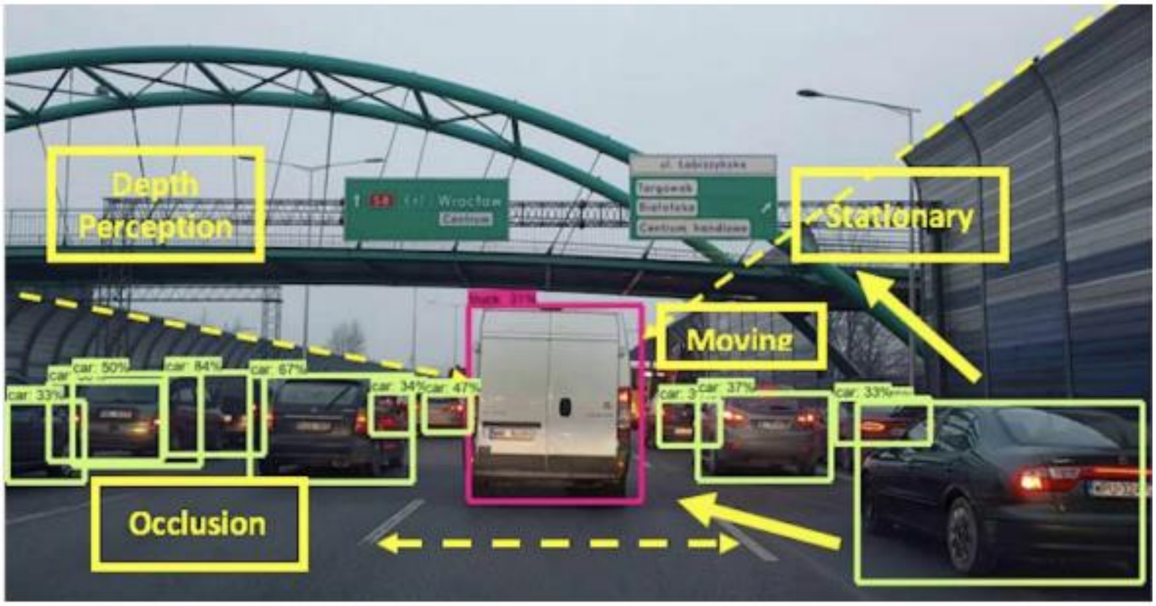

Figure 8: CSK-based object detection determining the location of objects and their properties [31].

Object detection is quite important for vehicles and needs to be conducted accurately or else disastrous consequences can occur. In 2016, a Tesla vehicle incorrectly identified a truck as an overpass and collided with it [30], [31]. Commonsense knowledge along with modern object detection systems such as YOLO [11] can be utilized to track not only the locations of objects, but also their properties as displayed in Figure 8. In the previously described situation, commonsense knowledge would determine and store the properties of the truck and know that it was not an overpass. Storing information can be used to determine future issues, such as rain indicating the road will be slippery since roads can remain slippery for a while after rain stops. In addition, autonomous vehicles benefit from storing objects even after they disappear from view since there are still nearby the vehicle. Commonsense knowledge can help autonomous vehicles avoid dangerous issues.

3. Chapter 3: Proposed Approach

3.1. Overview of Approach

Commonsense knowledge encompasses pragmatics, which relates to general world knowledge, and semantics, which relates to context-specific knowledge. Pragmatic knowledge is often useful for corner cases that do not occur regularly. For example, if the power goes out in a factory where a collaborative assembly robot is working, the robot should stop its current task until being given new tasks. Otherwise, it could run into a person or object while lighting is limited.

The main goal of this work is to determine how robots can support humans through utilizing commonsense knowledge while efficiently working. The two goals for robots collaborating with humans are as follows.

1. Determine commonsense priorities that can support and protect humans.
2. Determine commonsense priorities that can result in effective execution, especially in terms of minimal execution time.

These two goals must be balanced since if too much focus is placed on effective execution, humans may have worse work lives while if too little focus is placed on effective execution, production will be significantly harmed. Because of this, determining a balanced set of commonsense priorities is critical. Safety is the most important priority; robots must handle tasks leading to minimal human risk. Other priorities include the weight carried, the distance traveled and the danger and fragility of the carried parts. Lastly, the total execution time is quite important.

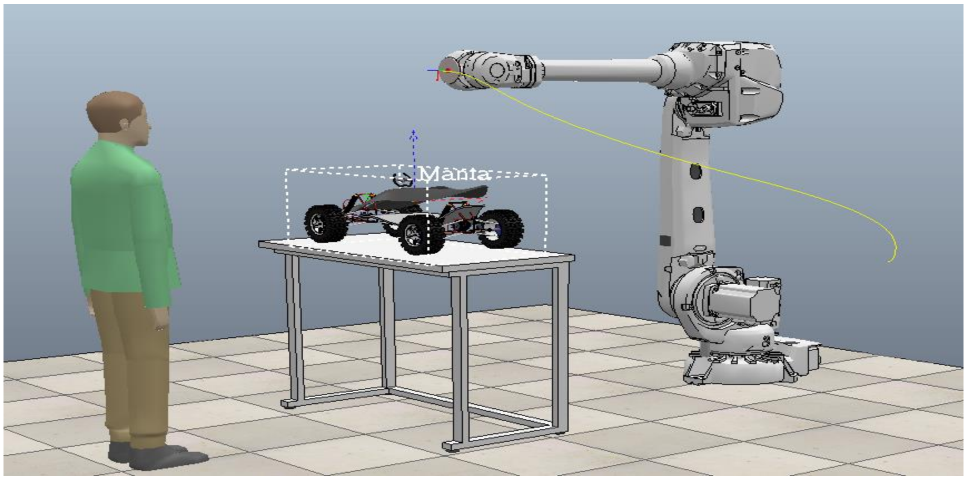

Figure 9: Example of HRC in Vehicle Manufacturing

Throughout this thesis, the term 'parts' will refer to the individual components used to create a final product by being combined in a pre-determined manner, while 'object' will refer to the assembled final product. These terms are used to explain how human robot collaboration for assembly occurs. For this human robot collaboration task, a human and a robot with one arm for grabbing objects will be cooperating to combine parts into an object. The robot uses commonsense knowledge in order to effectively select and move parts. The robot prioritizes heavier parts since humans can more easily and more quickly carry light parts. Humans will also have more difficulty with heavy parts and will therefore move more slowly. The robot arm prioritizes moving towards parts that are further away so work is easier for humans. The robot arm prioritizes carrying dangerous parts since that will keep humans safer. For example, the robot arm should carry sharp parts so humans do not cut themselves. At the same time the robot prioritizes carrying parts that are more stable since humans are better at handling parts that are fragile. The four main premises commonsense knowledge premises are as follows:

1. Humans prefer carrying lighter and closer parts due to ease
2. Humans will carry heavy parts more slowly than light parts.
3. Humans should handle less stable parts

4. Humans should not handle dangerous parts

Stability is defined as how likely a part is to remain stable if dropped; a wood part is more stable than a glass part. Humans will handle heavy parts more slowly than light parts, especially if carrying heavy parts all day. Humans will find it extremely frustrating if a robot mishandles an unstable part and not trust the robot to handle parts in the future. To avoid this scenario, humans will handle less stable parts. Since safety is the most important, robots will handle parts that are more dangerous.

The proposed approach for human robot collaboration is illustrated in Figure 8. There may be conflict between premises 3 and 4 occasionally. This framework for commonsense knowledge based human robot collaboration is defined as follows. First, human and robot execution affect a real world workspace. From there, task information is gathered from the workspace. That information is used to inform actor action analysis and a metric function, which are then combined into a cost function. The cost function in turn affects robot execution. Because of these conflicts and the need to determine how to best fulfill the premises, mathematical modelling is used.

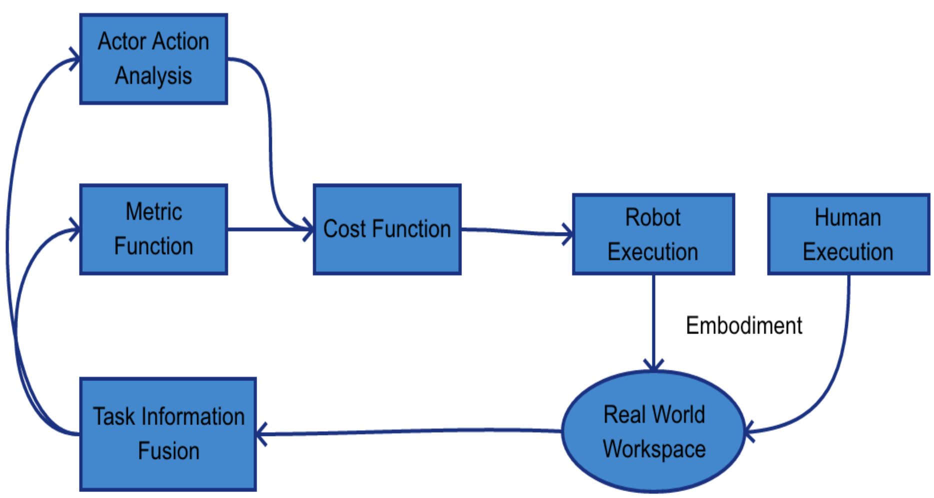

Figure 10: Framework of proposed approach in object assembly

The human robot collaboration system uses the CSK premises in order to optimize vehicle construction, with initial parts being shaped into a final object.

3.2. Details of Methodology

The knowledge base developed in this thesis focuses on human and robot priorities for selecting parts based on their properties, such as their weight, size, distance (the sum of the distance from the arm and the distance from their final position), danger and stability. In order to follow the premise that humans prefer lighter and closer parts, the robot arm should prioritize heavier parts and parts that are further away. In order to fulfill the premise that robots should not carry unstable parts, humans will carry parts that are less stable. Lastly, in order to actualize the premise that humans should not carry dangerous parts, robots should handle parts that are more dangerous.

The steps for arms handling parts are as follows.

1. Arms lock onto a part and indicate to other arms that they done so to prevent other arms form locking onto the same part.
2. Arms move to the part they have locked onto.
3. Arms move the part they are carrying to the final position.

Locking onto parts is used to indicate to other arms that a specified part is being targeted and that they should target other parts. From there, arms move to the parts they locked onto, grab them and then move to where they are supposed to be placed. This process is repeated until there are no remaining parts. An algorithm is used to define this behavior and is displayed below. The algorithm uses the real-time parts and their properties, and the arms' positions and priorities in order to determine which parts to first select and how to move them to their respective final positions.

Each arm uses a scoring algorithm based on its position and the attributes of the remaining parts to determine which part to select. The scoring equations are as follows.

$$W_a(s_p) = (min_a(s_p) + mean_a(s_p) + max_a(s_p))/3 \quad - (1)$$

$$O_1(p) = \sum_{a=1}^{t} r(a_{min}) \times W_a(s_p) /(a(p)) \quad - (2)$$

$$O_2(p) = \sum_{a=1}^{t} r(a_{max}) \times (a(p)/ W_a(s_p)) \quad - (3)$$

$$O(p) = O_1(p) + O_2(p) \quad - (4)$$

The set of parts is defined as s_p and the term a(p) refers to a specific attribute for the part p. $W_a(s_p)$ refers to the weighted average across the set of parts for a specific attribute, formed by averaging the set's minimum value for that attribute, maximum value for that attribute and the average value for that attribute. $R(a_{min})$ refers to an arm's priority for minimizing a specific attribute and $R(a_{max})$ refers to an arm's priority for maximizing a specific attribute. The series $\sum^{t}_{a=1}$ is the series of all measured attributes. There are currently four attributes being analyzed, but the system is customizable to permit adding other attributes. The relative attributes are important since the attribute of a part is compared against the weighted average for that attribute. The score is based on the priorities and values for attributes, including weight, danger and distance, so each attribute is finished.

For robot action planning, consider that humans and robots collaborate in the same workspace with the same set of parts denoted as *P*, and commonsense priorities denoted as *C*. Thus each *P* refers to a part such as wheel1, seat1 etc. while each *C* refers to a CSK priority such as distance, weight etc. Note that danger gets 1.5 times higher priority than the mean of all the other attributes. Here, *C(danger)* represents the CSK priority of this attribute while *C(x) for x = 1 to m* represents the priorities of each of the other attributes. Furthermore, P_i represents the i^{th} part in the set *P* while $S(p_i)$ represents the overall score for moving a part to its correct position. This score is calculated by comparing the attributes of a part against the maximum and minimum values of the attributes of all the parts, along with the CSK priorities. P_s represents the selected part while t represents the total number of parts. Thus, human and robot arms select their next part to move using Equations (5) to (7) herewith.

$$C(danger) = \frac{1.5}{m} \sum_{x=1}^{m} C(x) \qquad -(5)$$

$$S = \{S(p_i) \mid 0 \leq i \leq t-1\} \qquad -(6)$$

$$P_s = P(\text{argmax}(S)) \qquad -(7)$$

The following algorithms are used for task execution optimization in robot action planning and robot arm movement respectively.

Algorithm 1: Optimization of robot action planning

Input: Real-time parts, their properties, positions and priorities of arms
Output: The part that the current arm will target next

```
maxscore = 0; maxattrvals = [ ]; minattrvals = [ ];
selectedpart = None;
for (p = 0; p ≤ n; p++):
        for (a = 0; a ≤ t; a++):
                if (a(p) > maxattrvals [a]):
                        maxattrvals [a] = a(p)
                if( a(p) < minattrvals [a]):
                        minattrvals [a] = a(p)
for (p = 0; p ≤ n; p++):
        o(p) =0
        for (a = 0; a ≤ t; a++):
                o(p) += r(amax) ×  a(p) / maxattrvals [a]
                o(p) += r(amin) × minattrvals [a] / a(p)
        if (o(p) > maxscore):
                maxscore = o(p)
                selectedpart = p
return selectedpart
```

Algorithm 2: Optimization of robot arm movement

Input: Real-time parts, their properties, positions and priorities of arms

Output: The parts being placed in their correct positions

```
arm.holdingpart = false;
for arm in arms:
        if not (arm.lockedon):
                bestpart [arm] = arm.determinebestpart();
                bestconnect [arm] = arm.determinebestconnect();
                arm.lockonto (bestpart);
        if (arm.lockedon):
                if (not arm.holdingpart):
                        arm.movetopart(bestpart);
                        arm.pickup(bestpart);
                else:
                        arm.movetoconnect (bestconnect[arm]);
                arm.placepart (bestpart);
```

Within the simulation, human movement speed is affected by the amount of weight being carried, so the simulation reflects this with the following formula. In the simulation, a person can either be un-encumbered, slightly encumbered, encumbered, or very encumbered. The variable *E* is used to represent how encumbered a person is and the variable *v* represents their normal velocity: when the person is unencumbered. $V(E_s)$ represents the velocity when slightly encumbered, $V(E)$ represents the velocity when encumbered and $V(E_v)$ represents the velocity when very encumbered. This knowledge will modify their movement speed for placing parts using the equations as follows.

$$v(E_s) = 0.66 \times v \quad - (8)$$

$$v(E) = 0.5 \times v \quad - (9)$$

$$v(E_v) = 0.33 \times v \quad - (10)$$

These algorithms and equations are used for the execution of tasks in our proposed approach for human-robot collaboration. We now describe its experimental evaluation based on simulations.

4. Chapter 4: Experimental Evaluation

4.1. Experimental Platform

In order to determine optimal CSK priorities, the simulation was conducted with various robots priorities and with limitations placed on the robot priorities. Human priorities were determined by common sense and remained constant while robot priorities changed for different tests. The computer determined CSK priorities were tested against manually determined CSK priorities and simpler priorities. The CSK-based attributes corresponding to CSK priorities of distance, weight, danger and stability, and their ranges of values are depicted in TABLE I herewith, as coded in the KB and used in our experiments. The columns here indicate the CSK attributes. The first row defines the minimal human attribute values while the second row defines its maximal ones; the third row defines the maximal robot attribute values, and the fourth row defines its minimal ones.

	Distance (cm)	Weight (kg)	Danger (Level	Stability (Level)
Min for Human	5	1	0	0
Max for Human	600	55	30	30
Max for Robot	700	60	30	30
Min for Robot	10	2	0	0

TABLE I: COMPUTER DETERMINED PRIORITIES FOR HUMANS AND ROBOTS IN KB

The robot's maximization priorities are designed to mirror the human's minimization priorities and vice-versa. This allows humans to work with parts they prefer and are better at working with, resulting in object assembly being faster, safer and more effective. Robots can then handle those parts that humans have more difficulty moving, such as heavy, large or fragile parts.

The robot's priority ranges are limited in order to support human beings. This allows humans to work with parts they are best at working with, resulting in faster object assembly. Robots can handle heavy or large parts more effectively than humans, and therefore robots will handle those parts. Figure 11 shows an example of car parts in their initial and final state with respect to our simulation experiments.

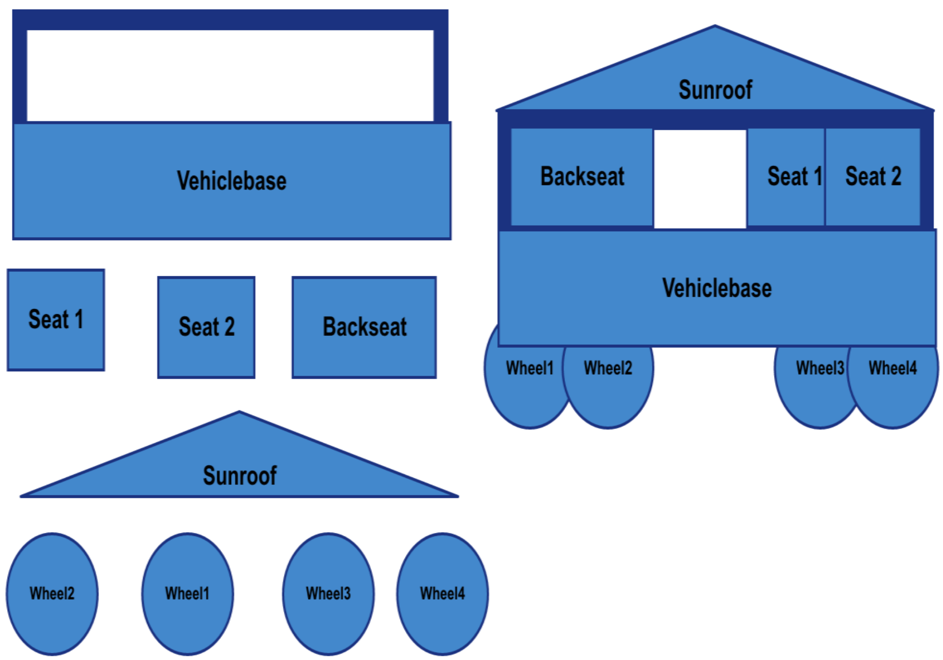

Figure 11: Car parts in their initial and final state

4.2. Task Description

The tasks consisted of simulated human robot collaboration, with humans and robots combining vehicle parts into a vehicle. In order to determine optimal CSK priorities, the simulation was run with various robots priorities and with positive and negative limitations placed on the robot priorities. Human priorities were determined by common sense and remained constant while robot priorities changed for different tests. The computer determined CSK priorities were tested against manually determined CSK priorities and simpler priorities. These simpler priorities include *blank*, where all priorities equal 0, *closest*, where all priorities except distance are equal to 0 and *norobot*, where only the human has priorities. When testing to determine the effectiveness of priorities, each set combined five different sets of parts 1000 times, with a total of 5000 executions. The attributes the human arm handled and the time were then average and stored. For the experiments, distance is measured in cm, weight in kg and time in seconds. Danger and stability are measured as relative levels. The exact combination of priorities is shown as follows in TABLE II. Note that we include the attribute "size" here since we initially defined it in the KB. However, we did not actually use the size of the parts within our simulation and in-person experiments so far. The other attributes, namely: distance, weight, danger and stability were used in the experiments conducted in this thesis. Addressing size of parts in the experiments remains an aspect of future work.

prioritiescsk

Distance	Weight	Danger	Stability	Size
0	0	0	100	0
100	100	50	0	100
100	100	150	100	100
0	0	0	0	0

prioritiescskv3

Distance	Weight	Danger	Stability	Size
0	0	0	50	0
0	250	100	0	0
100	150	100	0	0
0	0	0	0	0

blank

0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

closest

Distance	Weight	Danger	Stability	Size
0	0	0	0	0
100	0	0	0	0
0	0	0	0	0
100	0	0	0	0

norobot

Distance	Weight	Danger	Stability	Size
0	0	0	50	0
150	150	75	0	0
0	0	0	0	0
0	0	0	0	0

combinationscsk

Distance	Weight	Danger	Stability	Size
0	0	0	0	0
100	250	100	0	0
0	100	150	50	0
50	0	0	0	0

combinationscskv3

Distance	Weight	Danger	Stability	Size
0	0	0	0	0
100	250	100	50	0
0	250	150	0	0
100	0	0	0	0

thesispriorities

Distance	Weight	Danger	Stability	Size
0	0	0	0	0
100	100	150	50	0
0	50	200	50	0
100	0	0	0	0

TABLE II: LIST OF ALL PRIORITIES

In addition, the objects to assemble ranged from simple to complicated, allowing the simulation to test the priorities on disparate sets of parts. Through this, a more general set of priorities could be determined. TABLE III depicts a list of all sets of attributes in the experiments as well as the specific attributes for each set.

Attributesv1

partid	Label	isbase	Length	Width	Height	Weight	Danger	Stability
0	wheel1	FALSE	3	2	3	10	3	2
1	wheel2	FALSE	3	2	3	10	3	2
2	wheel3	FALSE	3	2	3	10	3	2
3	wheel4	FALSE	3	2	3	10	3	2
4	vehiclebase	TRUE	18	14	10	30	10	15
5	seat1	FALSE	4	4	5	5	2	4
6	seat2	FALSE	4	4	5	5	2	4
7	backseat	FALSE	12	4	5	15	5	8

Attributesv2

partid	Label	isbase	Length	Width	Height	Weight	Danger	Stability
0	wheel1	TRUE	6	3	6	2	3	4
1	wheel2	TRUE	6	3	6	2	3	4
2	wheel3	TRUE	6	3	6	2	3	4
3	wheel4	TRUE	6	3	6	2	3	4
4	tire1	FALSE	9	3	9	3	3	6
5	tire2	FALSE	9	3	9	3	3	6
6	tire3	FALSE	9	3	9	3	3	6
7	tire4	FALSE	9	3	9	3	3	6
8	frontmirror	FALSE	15	15	3	2	15	2
9	rearmirror	FALSE	15	15	3	2	15	2
10	vehiclebase	TRUE	54	42	30	15	10	15
11	vehiclebottombase	TRUE	54	42	30	15	10	15
12	seat1	FALSE	12	12	15	7	2	4
13	seat2	FALSE	12	12	15	7	2	4
14	backseat	FALSE	36	12	15	11	5	8
15	sunroof	FALSE	21	21	6	7	15	2

Attributesv3

partid	Label	isbase	Length	Width	Height	Weight	Danger	Stability
0	wheel1	FALSE	12	3	12	10	3	4
1	wheel2	FALSE	12	3	12	10	3	4
2	wheel3	FALSE	12	3	12	10	3	4
3	wheel4	FALSE	12	3	12	10	3	4
4	vehiclebase	TRUE	18	14	10	30	10	15
5	seat1	FALSE	12	12	15	5	2	4
6	seat2	FALSE	12	12	15	5	2	4
7	seat3	FALSE	12	12	15	5	2	4
8	front	FALSE	12	30	30	15	8	10
9	back	FALSE	20	30	30	15	8	10
10	roof	FALSE	48	30	12	20	12	4

Attributesv4

partid	Label	isbase	Length	Width	Height	Weight	Danger	Stability
0	wheel1	FALSE	8	4	8	5	3	4
1	wheel2	FALSE	8	4	8	5	3	4
2	wheel3	FALSE	8	4	8	5	3	4
3	wheel4	FALSE	8	4	8	5	3	4
4	wheel5	FALSE	8	4	8	5	3	4
5	wheel6	FALSE	8	4	8	5	3	4
6	wheel7	FALSE	8	4	8	5	3	4
7	wheel8	FALSE	8	4	8	5	3	4
8	vehiclebase	TRUE	50	14	10	30	10	15
9	seat1	FALSE	12	12	15	5	2	4
10	seat2	FALSE	12	12	15	5	2	4
11	seat3	FALSE	12	12	15	5	2	4
12	front	FALSE	12	30	30	15	8	10
13	back	FALSE	20	30	30	15	8	10
14	roof	FALSE	50	30	12	20	12	4

Attributesv5

partid	Label	isbase	Length	Width	Height	Weight	Danger	Stability
0	wheel1	FALSE	3	2	3	15	3	2
1	wheel2	FALSE	3	2	3	15	3	2
2	wheel3	FALSE	3	2	3	15	3	2
3	wheel4	FALSE	3	2	3	15	3	2
4	vehiclebase	TRUE	18	14	10	40	10	15
5	seat1	FALSE	4	4	5	10	2	4
6	seat2	FALSE	4	4	5	10	2	4
7	backseat	FALSE	12	4	5	20	5	8

TABLE III: LIST OF ALL SETS OF ATTRIBUTES AND ATTRIBUTES FOR EACH SET

In-person testing of human robot collaboration with commonsense knowledge has occurred as well. The CSK principles applied in the simulation were applied in the lab as well. For this lab, a model vehicle with four base parts and four wheels was used, as shown in Figure 12 below.

Figure 12: Model vehicle used for vehicle assembly experiments

The base parts are the cargo bed, the backseat, the front seat and the front, with the wheels being attached to the cargo bed and the front. A robot arm collaborates with a human to assemble the vehicle by grabbing the base parts and delivering them to the human. From there, the human attaches the wheels to the base parts. This division of labor is efficient since the robot cannot effectively handle the wheels while the human can attach the parts, and the robot can lessen the human's work by handling the base parts. The setup is shown below in Figure 13, and displays the experiments that occurred with a real robot arm.

Figure 13: Human-robot collaboration in vehicle assembly (snapshot of real lab demo)

4.3. Results Analysis

In the simulation experiments, the average attributes the human handled and the average time are summarized in the following table, i.e., TABLE IV.

attributes options	distance (cm)	weight (kg)	stability	danger	time (s)
csk	516.5	41.9	24.5	21.1	35.6
cskv3	561.2	45	28.2	22.9	35.5
blank	561.7	48	26.9	26	36.7
closest	502.8	48.3	26.4	25.6	37.6
norobot	486.6	54.8	29.9	30.8	38.6
work	495.2	39.8	24.9	19	35.2
workv3	502.9	39.3	24.8	19.9	35.3

TABLE IV: AVERAGED VALUES OF HUMAN ATTRIBUTES AFTER 5000 EXECUTIONS.

These results display benefits of using CSK priorities for HRC. For example, in the *csk* row in Table IV, the human carries less weight (compared to *blank*, *closest* or *norobot*), lessening the impact on human stamina. Stamina is not an issue if work is only done for a minimal amount of time, but if work is undertaken for hours at a time, stamina degrading will result in humans working slower. Having humans handle lighter parts will maintain stamina, and increase comfort as a bonus. Most importantly, danger is lowered with CSK priorities, especially in *combinationcsk* (and its *v3*), hence enhancing human safety. Execution time is also found to be lower with CSK priorities in comparison to simpler priorities. While the human travels a greater distance in *csk* (compared to *closest* for example), that is less important than safety. The human also performs tasks more slowly since the closer parts are occasionally heavier and this causes humans to move more slowly. The performance is more optimal for *combinationscsk* and *combinationscskv3* than for *prioritiescsk* and *prioritiescskv3* (as shown earlier in TABLE II), since the values were determined algorithmically rather than by hand. The *combinationscsk* priorities represent the newest research for this thesis. Further optimization can be performed in future work.

The in-person human robot collaboration proved to be beneficial as well. For the in-person experiment, robot and human handle parts in a predetermined and optimized order. Currently, all of the parts have a starting constant position, with the four base parts standing on the four corners of the white cardboard base in Figure 12, and the four wheels being nearby the human worker. The execution proceeds in the following order.

1. Robot hands over cargo bed of truck to human
2. Human attaches back wheel 1 and back wheel 2 to cargo bed
3. Robot hands over truck back seat to human
4. Robot hands over truck front seat to human
5. Robot hands over truck front to human
6. Human attaches front wheel 1 and front wheel 2 to truck front

Testing this assembly order showed that assembling the vehicle with aid from the robot makes the task easier than assembling it without assistance. While the task takes more time to complete with the robot's assistance, human stamina will remain higher for large-scale vehicle assembly execution, allowing humans to continue producing high efficiency and high quality work. Maintaining stamina becomes more relevant when in a large-scale setting, where a task is executed hundreds of times. Because of this, the human-robot collaboration outlined in these experiments is a significant contribution.

4.4. Discussion on Evaluation

The simulation results demonstrate that using commonsense knowledge for human robot collaboration makes work easier for humans while only slightly increasing the completion time in some cases. Humans are also safer since are carrying parts that are less dangerous on average. While work can be completed faster, oftentimes more danger is added, which can increase the chance of injury when tasks are frequently repeated. Having two humans collaborate to assemble the parts into a final object would be an option, but they would eventually become tired and work more slowly than a human and a robot collaborating. When tasks are repeated several times a day in a real scenario, avoiding tiredness is important, especially since it can help with preventing injury. Adding more aspects of commonsense knowledge can be even more effective than shown in the current simulation.

The lab work led to relevant inferences as well. Our work proves that humans and robots guided by CSK can be efficient in task execution while valuing human safety and comfort by protecting humans. The robot arm is also capable of verbally greeting the human worker and informing the human worker when it has brought a base part to them. This provides a sense of collaboration, which can be furthered by the human occasionally speaking to the robot. The robot arm made assembling the vehicle more efficient and effective. The lab experiments, even more so than the simulation studies, demonstrate task optimization in collaborative robotics, moving closer towards real executions for industrial vehicle assembly. This work contributes to smart manufacturing, in manner similar to other works of literature [6], [33], [34].

4.5. Subject Evaluation

There are no subjects for this study, due to the COVID-19 pandemic. The only experiments that have been conducted in this thesis are simulations or in-person experiments executed by the student and their committee advisors.

4.6. Limitation of Study

The study is limited due to the fact that the CSK system was only tested in the real world with one set of parts, where the parts' starting positions remained constant. However, results still appear conclusive that robots utilizing commonsense knowledge can help with improving human robot collaboration. The simulation also was not tested with complex sets of parts based on a real vehicle. Future work could remedy these issues.

5. Chapter 5: Conclusions and Future Work

5.1. Conclusions

This thesis surveys existing research on human robot collaboration, commonsense knowledge and autonomous vehicles, demonstrates the benefits and applications of human-robot collaboration, along with providing human-robot collaboration applications in manufacturing. The simulations display how human robot collaboration can be improved by applying commonsense knowledge, resulting in a better work environment for humans while retaining high efficiency. The in-person experiments display how the presented theory of applying commonsense knowledge to human robot collaboration is effective in practice. With the robot arm's assistance, assembling the vehicle has been made significantly easier. Applying human robot collaboration along with commonsense knowledge can help improve manufacturing.

In summary, the novel contributions of this thesis are as follows.

1. Providing a solution for vehicle assembly with CSK priorities to balance robot execution and aid humans in HRC
2. Mathematical modeling for robot action planning to provide task optimization
3. Conducting simulation tasks along with lab experiments to prove that humans and robots guided by CSK can be efficient in task execution while valuing human safety and comfort by protecting humans, making them carry lighter parts, making the collaborative experience very pleasant etc.

The work in the thesis has been published as research papers in the conferences IEEE IEMTRONICS 2020 (IEEE International IOT, Electronics and Mechatronics Conference) [35] and IEEE Big Data 2020 (IEEE International Conference on Big Data) [36], with the respective papers therein indicating different stages of the research. The paper on this work that appeared in IEEE IEMTRONICS received a best paper award in their Robotics track [37]. In addition, some part of the work has been submitted to a journal. The final outcomes of this thesis along with a detailed description on the approaches and experiments are in submission to another suitable journal.

5.2. Future Work

The simulation conducted in this thesis is modifiable, where more attributes can be added and modified based on the needs of the manufacturer. The system can be applied for larger and more complicated real tasks in the future. Currently, the arm used for in-person experiments does not detect the location of the parts; they are consistently placed in the same position. Future work can incorporate a detection system that would send the location of parts to the robot arm, which would then travel to the location and deliver the parts to a human worker.

Additionally, future work would consider the important concept of trust along with commonsense knowledge within the realm of human-robot collaboration, with priorities changing as trust increases for greater optimization. This would augment human-robot collaboration for a more enhanced experience in contexts such as smart manufacturing.

Furthermore, some experiments could be conducted in the future for subjective evaluation in real world human-robot collaboration. For example, this could consider factors such as the experience with in-person experiments being pleasant due to the conversation between the human and the robot. Other subjective evaluations could involve the manufacturing outcomes with respect to their reception by the real world. Some of this future work could potentially entail contacting domain experts from the industry in smart manufacturing. Their inputs on real world experiments and feedback through surveys etc. would be valuable in further stages of the work emerging from this thesis, on a larger scale.

In short, the future work emerging from this thesis is as follows.

- Adding more attributes in the task execution.
- Testing for larger and more complicated real world tasks.
- Having the robot arm detect part positions.
- Undertaking subjective evaluation experiments for real world human-robot collaboration.
- Contacting people from the smart manufacturing industry.
- Considering the concept of trust in human-robot collaboration, with priorities changing as trust increases for greater optimization.

In general, this thesis deploys concepts from commonsense knowledge, proposes an approach based on that for human-robot collaboration and executes the approach in the application of vehicle assembly within the contexts of smart manufacturing. Future work would provide further enhancements from all these perspectives, thereby making even stronger impacts on robotics and artificial intelligence.

BIBLIOGRAPHY

[1] J. Suchan and M. Bhatt, “Commonsense Scene Semantics for Cognitive Robotics: Towards Grounding Embodied Visuo-Locomotive Interactions,” 2017, doi: 10.1109/ICCVW.2017.93.

[2] Y. Chen *et al.*, “A Robotic Lift Assister: A Smart Companion for Heavy Payload Transport and Manipulation in Automotive Assembly,” *IEEE Robot. Autom. Mag.*, vol. 25, no. 2, pp. 107–119, 2018.

[3] D. E. Whitney, C. A. Lozinski, and J. M. Rourke, “Industrial robot forward calibration method and results,” *J. Dyn. Syst. Meas. Control. Trans. ASME*, vol. 108, no. 1, pp. 1–8, 1986, doi: 10.1115/1.3143737.

[4] T. Pettersen, J. Pretlove, C. Skourup, T. Engedal, and T. Lokstad, “Augmented reality for programming industrial robots,” in *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.*, 2003, pp. 319–320.

[5] J. Iqbal, R. U. Islam, S. Z. Abbas, A. A. Khan, and S. A. Ajwad, “Automating industrial tasks through mechatronic systems–A review of robotics in industrial perspective,” *Teh. Vjesn.*, vol. 23, no. 3, pp. 917–924, 2016.

[6] K.-D. Thoben, S. Wiesner, and T. Wuest, “‘Industrie 4.0’ and smart manufacturing-a review of research issues and application examples,” *Int. J. Autom. Technol.*, vol. 11, no. 1, pp. 4–16, 2017.

[7] B. Hayes and B. Scassellati, “Autonomously constructing hierarchical task networks for planning and human-robot collaboration,” 2016, doi: 10.1109/ICRA.2016.7487760.

[8] N. Tandon, A. S. Varde, and G. de Melo, “Commonsense Knowledge in Machine Intelligence,” *ACM SIGMOD Rec.*, 2018, doi: 10.1145/3186549.3186562.

[9] N. Tandon, G. De Melo, and G. Weikum, “WebChild 2.0: Fine-grained commonsense knowledge distillation,” 2017, doi: 10.18653/v1/P17-4020.

[10] E. Davis and G. Marcus, “Commonsense Reasoning and Commonsense Knowledge in Artificial Intelligence.” [Online]. Available: http://cs.nyu.edu/faculty/davise/papers/WS.html.

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” 2016, doi: 10.1109/CVPR.2016.91.

[12] A. Kilgarriff and C. Fellbaum, “WordNet: An Electronic Lexical Database,” *Language (Baltim).*, 2000, doi: 10.2307/417141.

[13] F. M. Suchanek, G. Kasneci, and G. Weikum, “Yago: A Core of Semantic Knowledge,” 2006.

[14] J. Lehmann *et al.*, “DBpedia - A large-scale, multilingual knowledge base extracted from Wikipedia,” *Semant. Web*, vol. 6, no. 2, pp. 167–195, 2015, doi: 10.3233/SW-140134.

[15] H. Liu and P. Singh, “ConceptNet - a practical commonsense reasoning tool-kit,” *BT Technol. J.*, vol. 22, no. 4, pp. 211–226, 2004, doi: 10.1023/B:BTTJ.0000047600.45421.6d.

[16] “Quartz Weekly Obsession: Google Translate,” *Quartz*, 2020. https://qz.com/emails/quartz-obsession/1781599/.

[17] A. Varghese, A. S. Varde, J. Peng, and E. Fitzpatrick, “A Framework for Collocation Error Correction in

Web Pages and Text Documents."

[18] S. Pellegrinelli, A. Orlandini, N. Pedrocchi, A. Umbrico, and T. Tolio, "Motion planning and scheduling for human and industrial-robot collaboration," *CIRP Ann. - Manuf. Technol.*, 2017, doi: 10.1016/j.cirp.2017.04.095.

[19] B. Mutlu, A. Terrell, and C. Huang, "Coordination Mechanisms in Human-Robot Collaboration," 2013.

[20] J. Richer and J. L. Drury, "A video game-based framework for analyzing human-robot interaction: characterizing interface design in real-time interactive multimedia applications," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on human-robot interaction*, 2006, pp. 266–273.

[21] W. Wang, Y. Chen, and Y. Jia, "Which way is the best?: Evaluation and optimization of dual-arm robot path planning for human-robot collaborative tasks in smart manufacturing contexts," *ASME 2019 Dyn. Syst. Control Conf. DSCC 2019*, vol. 1, no. 1, 2019, doi: 10.1115/DSCC2019-9143.

[22] Y. Guo and L. E. Parker, "A distributed and optimal motion planning approach for multiple mobile robots," in *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*, 2002, vol. 3, pp. 2612–2619.

[23] E. Matheson, R. Minto, E. G. G. Zampieri, M. Faccio, and G. Rosati, "Human-robot collaboration in manufacturing applications: A review," *Robotics*. 2019, doi: 10.3390/robotics8040100.

[24] M. Chen, S. Nikolaidis, H. Soh, D. Hsu, and S. Srinivasa, "Planning with Trust for Human-Robot Collaboration," 2018, doi: 10.1145/3171221.3171264.

[25] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Rob. Auton. Syst.*, 2009, doi: 10.1016/j.robot.2008.10.024.

[26] J. P. Vasconez, G. A. Kantor, and F. A. Auat Cheein, "Human–robot interaction in agriculture: A survey and current challenges," *Biosystems Engineering*. 2019, doi: 10.1016/j.biosystemseng.2018.12.005.

[27] "SAE International Releases Updated Visual Chart for Its 'Levels of Driving Automation' Standard for Self-Driving Vehicles," *SAE*, 2018. https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-"levels-of-driving-automation"-standard-for-self-driving-vehicles.

[28] R. Beard *et al.*, "Autonomous vehicle technologies for small fixed-wing UAVs," *J. Aerosp. Comput. Inf. Commun.*, vol. 2, no. 1, pp. 92–108, 2005, doi: 10.2514/1.8371.

[29] "On the Road to Full Autonomy: Self-Driving Cars Will Rely on AI and Innovative Memory," *Micron Lettered in White*. .

[30] A. Pandey, M. Puri, and A. Varde, "Object detection with neural models, deep learning and common sense to aid smart mobility," 2018, doi: 10.1109/ICTAI.2018.00134.

[31] P. Persaud, A. S. Varde, and S. Robila, "Enhancing autonomous vehicles with commonsense: Smart mobility in smart cities," 2018, doi: 10.1109/ICTAI.2017.00155.

[32] U. E. Manawadu, M. Kamezaki, M. Ishikawa, T. Kawano, and S. Sugano, "A hand gesture based driver-vehicle interface to control lateral and longitudinal motions of an autonomous vehicle," in *2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 - Conference Proceedings*, 2017, pp. 1785–1790, doi: 10.1109/SMC.2016.7844497.

[33] F. Tao, Q. Qi, A. Liu, and A. Kusiak, "Data-driven smart manufacturing," *J. Manuf. Syst.*, 2018, doi: 10.1016/j.jmsy.2018.01.006.

[34] P. O’Donovan, K. Leahy, K. Bruton, and D. T. J. O’Sullivan, “An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities,” *J. Big Data*, 2015, doi: 10.1186/s40537-015-0034-z.

[35] C.J. Conti, A.S. Varde and W. Wang, "Robot Action Planning by Commonsense Knowledge in Human-Robot Collaborative Tasks", IEEE IEMTRONICS, 2020, doi: 10.1109/IEMTRONICS51293.2020.9216410.

[36] C.J. Conti, A.S. Varde and W. Wang, "Task Quality Optimization in Collaborative Robotics", IEEE Big Data conference, 2020.

[37] Computer Science Student and Faculty Receive Best Paper Award at IEEE IEMTRONICS 2020, https://www.montclair.edu/csam/2020/09/14/computer-science-student-and-faculty-receive-best-paper-award-at-ieee-iemtronics-2020/

APPENDIX A. THE NOUN PROJECT

URL - https://thenounproject.com/

Man by Viktor Vorobyev from the Noun Project

Robot by iconsmind.com from the Noun Project

Cube (Part) by Noe Araujo from the Noun Project

Glass by Gregor Cresnar from the Noun Project

Can by S. Salinas from the Noun Project

Bottle by AFY Studio from the Noun Project