中央大学博士論文

Responsible Robots
– A Novel Approach to Safe and Productive Human-Robot Collaboration

Audun Rønning Sanderud

博士（工学）

中央大学大学院
理 工 学 研 究 科
精 密 工 学 専 攻

平成２７年度
２０１６年3月

Summary

With the more open innovation model seen in the later years, small and medium enterprises have a growing importance in the industry. These types of companies require robotic equipment that is highly flexible, but also easy to use. An important approach to simple and flexible use of robots is through human-robot collaboration. In a human-robot collaboration, one can combine the strengths of the co-workers, the strength and repeatability of the robot, and the flexibility and adaptability of the human. However, there are still many challenges in the way before a high level of safe and productive human-robot collaboration can be fully realized. One of the most critical challenges is the safety issue. If the human is to work along side a robot, a system must be able to ensure the human operator's safety.

However, is it enough to be merely safe? As the human and robot co-workers' collaboration grow closer, the importance of the human's aspect of the collaboration grows and a more advanced robot co-worker is required. Is it a selling point for a human employee that he is safe to work with? The safety strategy for robots has not changed much in the past decades. Several approaches with the basic strategy of moving away if the robot is too close to the human operator have been proposed. Systems like these are also needed to realize safe human-robot collaboration, but again, is it enough that the robot is safe? Moreover, even when these safety systems works properly, it is not avoiding human-robot conflicts, they simply react when a danger is imminent. These conflicts are disturbing for the human operator and interrupts his/her concentration. Furthermore, the robot is not even able to complete its task if it is forced to avoid the human. A new safety strategy for safe and productive human-robot collaboration is therefore needed. This system should be proactive against dangers and aspire to maintain the productivity of the system. In this way, the human operator and the robot should be able to harmonize and improve their work together.

This thesis proposes a novel strategy for safe and productive HRC called Responsible Robots. A Responsible Robot is a robot sharing the responsibility for the productivity and the safety in the collaboration. While it has previously been the full responsibility of the human to set proper safety rules for the robot, this should be a joint venture. The Responsible Robot acts proactively against dangers and it can in this way plan when to execute its different tasks to ensure the safety of the human operator while being productive. A model for realizing Responsible Robots is then proposed.

The model to realize Responsible Robots enhances the system's situation awareness by adopting a risk perception. The system observes the human operator and learns from his/her work patterns. The risk perception enables the system to estimate the risk associated with each of the robot's tasks. The system can then select the task with the lowest risk and postpone high risk task in case the risk is reduced later in the operation. This way, the system acts proactively against dangers and may reduce the number of human robot conflicts. The system can plan its tasks better and keep up the productivity to a greater extent than a pure reactive safety system. The reduced number of human-robot conflicts can also have a positive effect on the human operator, as he/she will not be disturbed as often as before.

The proposed model was implemented in an experimental setup and tested with several human test subjects. The experimental setup was realized with a heavy-duty NACHI MR20 7-axes industrial robot. The experiments demonstrated the system's ability to make safe proactive decisions. It clearly reduced the number of human robot conflicts. Further, the experiments showed that the system was able to maintain its productivity while being safe. Lastly, the effect the system had on the human operator was tested. The experiments showed that the system was able to reduce the workload for the human, and there were also indications that the system reduced the stress level for the human. It was therefore concluded both that the proposed system fulfills the requirements of a Responsible Robot, and also that a HRC could be improved by implementing Responsible Robots.

Preface

This PhD thesis documents a research project carried out at Chuo University. The project has been carried out in collaboration between the Human-System Laboratory at Chuo University and PPM AS. The project was initiated in September 2012 and ended in August 2015.

The work is primarily intended for research on human-robot collaboration with industrial robots. It is my hope that my ideas and findings can be of interest also to other branches of robotics and that Responsible Robots will inspire a more comprehensive way of thinking about safety in robotics.

Audun Rønning Sanderud
Trondheim, January 2016

Contents

Nomenclature

Abbreviations

BSS	Brier Skill Score
CDF	Cumulative Distribution Function
DBS	Decomposed Brier Score
FP	False Positives
FPR	False Positive Rate
HHC	Human-Human Collaboration
HRC	Human Robot Collaboration
HSW	Human Sub-Workspace
LMA	Levenberg-Marquardt Algorithm
MSE	Mean Square Error
PDF	Probability Density Function
PP-HRC	Preprogrammed Human-Robot Collaboration
PS	Problem Statement
REL	Reliability (Part of the DBS)
RES	Resolution (Part of the DBS)
RR-HRC	Responsible Robots based Human-Robot Collaboration
RRM	Risk Reducing Measure
RSW	Robot Sub-Workspace
SA	Situation Awareness
SMEs	Small and Medium Enterprises
SSW	Shared Sub-Workspace
TBV	Time Between Visits
ToV	Time of Visit
TP	True Positives
UNC	Uncertainty (Part of the DBS)

Parameters

α	Scaling multiplier
η	Parameter in Human Motion Prediction
γ	Design Parameter Danger Field

λ	Scale Parameter
μ	Location Parameter
ρ	Maximum distance of proximity field
σ	Shape Parameter
k_1	Design Parameter Danger Field
k_2	Design Parameter Danger Field
L	Limb type factor
l	Length of human velocity prediction history

Variables

Δt_0	Current time since last visit
Δt_V	Time between two visits
δ	Distance from tracked point to point in proximity field
κ	Proximity Factor
$\mathbf{s}$	Observed object's point in space
$\mathbf{s}_t$	Arbitrary point in space
$\mathbf{v}_t$	Next iterations predicted velocity
ε	Fitness improvement threshold
ADV	Average Daily Visits
c	Consequence
f	Fitness
G	Maximum number of PDFs
p_i	Probability that voxel i is selected for refitting
T	Task execution time of the robot task
t	Time
t_0	Current time
t_V	Time of a visit
v_t	Limb Velocity
g	Available number of PDFs

Chapter 1

Introduction and Problem Formulation

1.1 Introduction

With the more open innovation model seen in the later years [1], Small and Medium Enterprises (SMEs) have a growing importance in the industry. These types of companies require robotic equipment that is highly flexible, but also easy to use, both with regards to programming and operation. The most important approach to simple and flexible use of robots is through Human-Robot Collaboration (HRC)(Figure 1.1) and redundant robotics [2]. An HRC system must be designed to fit the level of collaboration, but at any level the system should allow the operator to focus fully on his or her task, and not be concerned with where the robot is, or its current task. The robot should autonomously give the best assistance and avoid collisions at all times. A reliable safety strategy is therefore vital.

The most widespread protection strategy practiced in the industry today is based on isolating robots from their surrounding environments [3]. While some HRC systems are commercially available, they have some major limitations.

Rethink Robotics™ have developed the Baxter system[1]. Baxter is a double seven-axes arm, with a fully integrated control system. It can be installed in one hour and does not require any safety installations beyond the built-in safety system. But with only 2.3 kg payload per arm the work is limited to very light operations.

ABB have introduced the SafeMove system which is designed to bring the operator closer to the industrial robot [4]. SafeMove operate with zones in which the operator can move safely, and allows a more efficient use of the robot. The robot will automatically slow down as the operator approaches, and go to a full stop if the operator is too close.

The SMErobots™ initiative has done extensive research on, and developed systems to simplify both the industrial level programming and safety issues related to industrial robot installations[2].

[1]Rethink Robotics™, www.rethinkrobotics.com (Accessed 11/10/2015)

[2]SMErobot ™, The european robot initiative for strengthening the comprehensiveness of SMEs in manufacturing, www.smerobot.org (Accessed 11/10/2015)

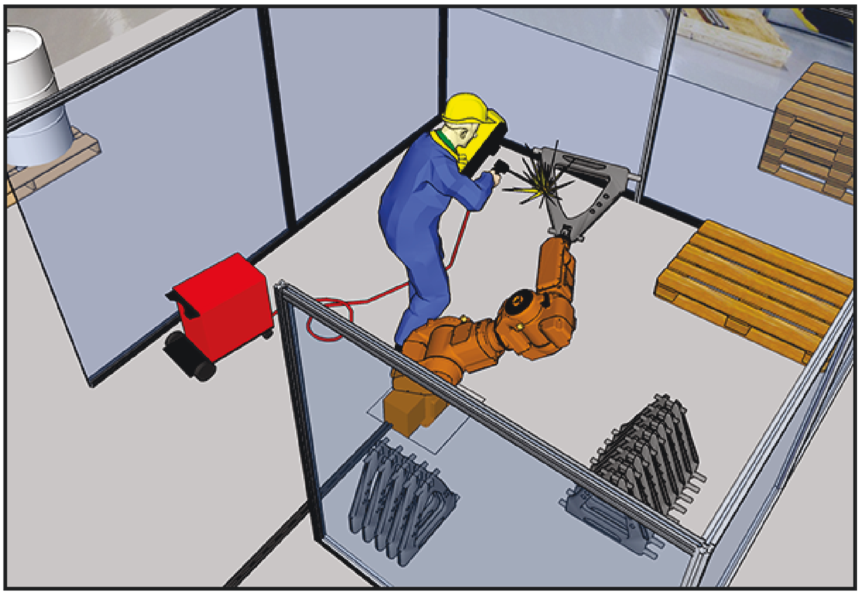


Figure 1.1: A collaboration between a human and a robot on a welding operation.

MRK Systeme have a commercially available safety system for selected low payload robots[3]. The system includes a capacitive cover for the robot, ensuring a full stop if the operator comes in contact with it.

Moreover, the current safety standards allow very limited human-robot collaboration [3]. The systems for HRC available in the industry today are naturally limited by this. Strict regulations on maximum allowable payload, speed etc makes the aforementioned system that is possible within today's standard [5], [6].

A new standard covering collaborative robots is currently under development [7]. The standard will allow closer human-robot collaboration, given that a set of performance control methods are implemented. Separation monitoring is one of these performance control methods. This is a system which at all times ensures that the robot manipulator is at a certain distance from the human operator to avoid injuries. Such a system will require an advanced sensor system and algorithms to reconfigure the robot manipulator based on sensor readings. With separation monitoring in place, a company can do highly complex human-robot collaborative tasks, reducing some of the pressure on programming of the robot prior to the operation. Also larger enterprises, such as the car manufacturing industry, can benefit from an efficient separation monitoring system. Just imagine a production line with tens of robots, and an operator walking among them, supervising and making adjustment while in full operation. The challenge lies in achieving this without compromising the productivity of the robot.

However, is a separation monitoring system enough to achieve a fruitful collaboration between a human and a robot? Is it enough that the robot is merely safe? The three laws of robotics were formulated by Isaac Asimov in his short story "Runaround" from 1942 [8]. Although the work is purely fictional, the laws are often cited and referred to as a guiding principal in HRC design. The three laws appear in many alterations by Asimov and other authors, however, the original laws are:

[3]MRK-systeme GMBH, www.mrk-systeme.de (Accessed 11/10/2015)

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given to it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

The laws suggest a robot that actively prevents a human to come to harm, not just passively obey the safety rules set by the human. The robot should even deny an order given to it by the human if it might cause a hazard. It is important not to get too caught up in the laws as they are a work of fiction. However, the active attitude towards safety is an intriguing aspect of HRC.

1.2 Problem Formulation

The common denominator of the majority of the approaches to solve the safety issue in HRC is the robot's reactive response to danger. With these approaches, the robot becomes a passive contributor towards safety as it only react on rules and commands set by the human integrator, such as keeping a minimum separation, or reduce the speed when in proximity to the human operator. This passive approach reduces the robot's ability to plan and make task related decisions, as it has no means of including safety related issues in its decisions. Therefore, the entire safety responsibility lies with the human integrator and human operator. This responsibility is an added workload and stress factor for the human operator. A goal for safe HRC should be robots that actively make decisions not to harm the human being. A robotic system that decides its action also on the basis of safety, and actively avoids actions that could harm the human, would take some of the safety related responsibility in the collaboration. The robot would share the responsibility with the human operator, relieve some of the operator's pressure, and in this way allow for more focus on the task.

1.3 Structure of the Report

This thesis is mainly structured around three parts. The first part includes the introductory Chapter 1 and the related work described in Chapter 2. Chapter 1 introduces the topic and defines the problem formulation for this research. Related work on this topic is investigated and compared in Chapter 2.

In the second part of this thesis, the main contributions to this research is presented. Chapter 3 discusses how human beings make decisions, and how this can be used to enhance the robot's decision making with respect to safety. A model that uses a risk analysis framework to realize this enhancement is then presented. Chapter 4 presents and describes the necessary components in the presented model and how this model can benefit many aspects of robotics.

The third and last main part of the theses consists of the three chapters 5-7. The three chapters present the experimental tests that were conducted. The performance of the components related to the likelihood analysis are investigated in Chapter 5. In Chapter 6 the performance of the system is tested with several test subjects. The effects on the perceived workload for the operator is investigated in Chapter 7. The three chapters with the experiments are written to be independent to some extent. Components in the experimental setup that are the same in two or more of the experiments are described in each of the relevant chapters.

The final chapter discusses the proposed new strategy and novel model to realize this strategy. The results from the experiments are summarized and a conclusion is drawn along with some suggestions for further work.

Chapter 2

Related Work

2.1 Introduction

For many years, researchers have worked towards humans and robots working alongside each other as coworkers. Although many levels of collaboration exists, it is generally agreed that a human-robot collaboration is whenever a human and a robot is working on the same or separate tasks in the same shared work space. There is still need for more research before a high level of HRC can be realized. To understand and discover which part of the HRC puzzle is missing, it is necessary to study the existing approaches in literature on the topic and related topics.

In this research, it is distinguished between approaches that are commercially available, and approaches in research. Approaches that are commercially available for the industry must comply with ISO 10218 [5], [6] and are thus fairly limited by that. Furthermore, the approaches presented in research are generally focusing on the task understanding aspect of the robot, or the safety aspect. Today's approaches towards safe HRC are predominantly inspired by the upcoming standard on collaborative robots [7]. The new standard allow closer collaborations as long as the system holds certain performance criteria. One of which is the separation monitoring previously mentioned. Separation monitoring is a system that ensures a minimum separation between the human and the robot at all times. Moreover, this research distinguish between approaches that are proactive and reactive, both for task and safety based approaches. A reactive approach acts only when certain input activates it and can be easily associated with a feedback loop in control theory. Conversely, a proactive system can be associated with a feed-forward loop and is designed to take action before it reach the critical input. From the human operator's perspective, this difference in behavior is more important than how the actual reactive or proactive behavior is realized in control.

While safe HRC with industrial robots is the focus of this research, many interesting approaches have also been presented in related research fields. Some of these might have some relevance and applicability for industrial robots. Therefore, some approaches to HRC with mobile robots will be presented last in this chapter. Firstly, relevant research on other aspects of HRC is presented. This include research on the effects of working alongside a robot has for the human operator, and the importance of proper team work models. The section will also include some research on HRC that is not directly related to improving the robotic system's functionality and performance.

At the end of the chapter, the most important findings will be highlighted and a series of problem statements will be formulated. Some criteria for fulfilling these statements will also be discussed. These statements will serve as the main guide and road map throughout the thesis.

2.2 Miscellaneous Research on HRC

2.2.1 Introduction

Research and advances in HRC is predominantly related to the performance of the robot. However, there are many other aspects that should be considered, such as how different levels and types of collaboration need special care in the design of the collaborative work cell. This section will present some of the current research on in this field.

2.2.2 Miscellaneous Human-Robot Collaboration Research

As aforementioned there are many levels of collaboration with a robot, much like with a human. Four degrees of interactions between humans and robots were identified in a study by Helms et al. [9], [10]: independent, synchronous, simultaneous, and supportive. In independent work, the human and the robot operate independently on different work pieces. This is the case which is most similar to today's industrial robots. The human operator and robot work consecutively on the same work piece in a synchronous collaboration. In simultaneous work the human operator and robot work on the same work piece, however physically separated. The closest collaboration occurs when the human operator and robot is collaborating on the same task on the same work piece. E.g. in a grinding operation the robot could carry the weight of the grinding tool and follow a rough path, while the human adjusts the path along the way by manipulating the grinding tool through force control directly in the operation. This is often also referred to as a human-in-the-loop approach.

How HRC best can be used, and which challenges HRC poses have also been investigated [11], [12]. Application of HRC in the industry is shown to extend the applicability of industrial robots to a larger part of industrial production.

A risk assessment for HRC was presented by Marvel et al. [13]. The assessment gives a thorough description of hazards and severity associated with HRC. By decomposing the task, the safety can be evaluated based on the subtask and proper risk mitigating actions can be selected. The assessment was applied in several case studies, both pre and post implementation of risk abatement. A reoccurring risk that did not have an efficient risk abatement was impacts. While this risk analysis is also very important, it is only a part of the design of the work cell and does not improve the performance of the robot during operation.

Marato et al. presented a system for HRC based on asynchronous task patterns. The design of the workcell allows the human and the robot to alternate who is working in a given subworkspace [14]. E.g while the human is retrieving parts to bring to the workspace, the robot is collecting products form the assembly area. This was realized through planning, thus rely on careful investigations prior to execution.

A study about new opportunities presented by HRC and the upcoming standard ISO 15066 [7] has been presented by Eder et al. [15]. Their results emphasizes the importance of safety and trust. Moreover, it is stated that the robotic co-worker must meet the innate expectations of the humans it work with. It is also important that it is able to communicate its attention to the human operator.

There are several examples of research on communication between the human operator and the robot. One of the approaches investigate use of speech, head poses and gestures as part of that communication [16]. This multimodal human-robot interaction is important to convey subtle information that might get lost in other more sterile forms of communication. An approach attempting to implement emotions in the robotic system has even been presented [17]. The system was implemented in a collaborative assembly task where the robots emotion changed on the basis of the task situation. The robots emotion was represented by a screen image and was implemented on the Baxter robot. Their results showed that a static emotion produces better assembly performance than that produced for the no emotion condition.

2.2.3 Summary

The importance and the impact HRC will have on the industry is well established. Several frameworks and categorizations exists for supporting the work with realizing HRC. However, there are many challenges yet to be solved. The importance of the robotic system being aware of the humans expectations of it have been shown. Also more complex human robot communication is needed for a smooth and safe collaboration.

2.3 Human-Robot Teams and Trust in Automation

2.3.1 Introduction

It is easy to mainly focus on the robot and forget the human when developing a system for HRC. However, it is clear that the both parties' contribution is equally important in a collaboration. Several researchers are devoted to investigate how humans and robots best can work is a team and what effects this nontraditional collaborator have on the human. The importance of this research grows as the collaboration between the human operator and the robot grows closer and a higher level of collaboration is reached.

2.3.2 Human Robot Teams

A study have been conducted by Idaho National Laboratory on the effects of sharing task execution responsibilities with the robot [18], [19]. The experiments investigated a middle ground between direct human control and full robotic autonomy. The human and a mixed-initiative mobile robot worked on various search and explore tasks. The experiments showed a positive effect from sharing the responsibility for the task, as opposed to each party only being responsible for their own task. Without loss of productivity, the humans felt reduced workload and fewer instances of confusion.

Keebler et al. with the Team Performance Laboratory at the University of Central Florida have done extensive research on applying well known human-human team heuristics to HRC applications [20], [21]. A "Wizard-of-Oz" approach was used to investigate how humans collaborate with robots in different team settings. The aim of the research is to develop proper team heuristics for HRC, from a human psychology perspective. Further, a proper framework for which heuristics to apply to different types of HRC is investigated. Nikolaidis et al. [22] have also applied cross training to human robot teams. This is a well known human-human team tool where every member of the team learn every other team member's tasks. This is beneficial for understanding how to improve your own task so that other team members might benefit from it.

Goodrich et al. propose an approach for a human to manage multiple robots [23]. They investigate use of two management styles and show their effect in an extensive experiment. It is shown that the level of autonomy is important to achieve the best team performance, and that adjusting quality parameters may give a greater advantage than adjusting the robots behavior.

A study of the level of autonomy in the team was conducted by Marble et al. [24], [25]. The study showed a great variation in novice users general ability to trust autonomous robots. A robot with an adjustable level of autonomy was used in an experiment. The experiment showed that the most experienced users completed their tasks faster with less autonomy, while inexperienced users needed more autonomy to complete the task. However, every participant were able to complete the task satisfactory. A robot that is able to detect and adjust its autonomy according to the user to achieve best possible task result would therefore be optimal.

2.3.3 Trust in Automation

Trust in automation have been studied by Hoffman et al. [26], [27], and found to be closely related to interpersonal trust. Interpersonal trust has been defined as a trustor's willingness to be vulnerable to a trustee's actions based on the expectations that the trustee will perform a particular action that is important to the trustor [28]. Further, research shows that interpersonal trust depends on several factors including perceived competence, benevolence, understandability, and directability. That is how rapidly the trustor can assert control if things go wrong. These factors are also important to trust in automation, along with some other factors. These include the technology's limitations and weaknesses, such as reliability, validity, utility, robustness and false-alarm rate [29], [30]. There is no doubt about the importance of trust in automation in HRC.

How interpersonal touch can affect the relationship between two humans is well established. It has for example been shown that waitresses who touch customers get more tip than waitresses that do not touch the customers [31]. Fukuda et al. [32] demonstrated the same effect in human-robot interaction. They suggest that touch can be used as a powerful communication channel in human-robot interaction. Although this might not be applicable in a industrial HRC, and demonstrates how powerful and applicable human-like attributes can be to build trust.

In a study by Short et al. [33] participants played games of "rock-paper-scissor" with a robot. The study aimed to investigate a humans social engagement while playing with a robot that was conditioned to cheat. It was shown that the participants playing with the cheating robot was much more engaged and active in the collaboration. While it would not be recommended to introduce cheating to industrial robots, the idea of adding human attributes to gain trust and keep the human focused and interested is an important contribution. Using a robot as a motivator in physical and mental task was investigated by Fasola and Matarić [34]. Human participants played a simple game while being motivated to continue by a robot. Their preliminary results demonstrated a positive effect on the human participants.

Summary

It is apparent that there are many important aspects to consider when designing a HRC work cell. There is no doubt that the effect the collaboration have on the human is important and should be considered when proposing new systems for HRC. The importance of trust in automation have been presented, and one of the factors to again this trust is shown to be the false-alarm rate. In other words, disturbing the human operator with unnecessary alarms unless they are strictly needed should be avoided. All in all, there are many examples in the literature that demonstrate a positive effect by introducing human-like attributes in the robotic system.

2.4 Commercially Available Approaches to HRC

2.4.1 Introduction

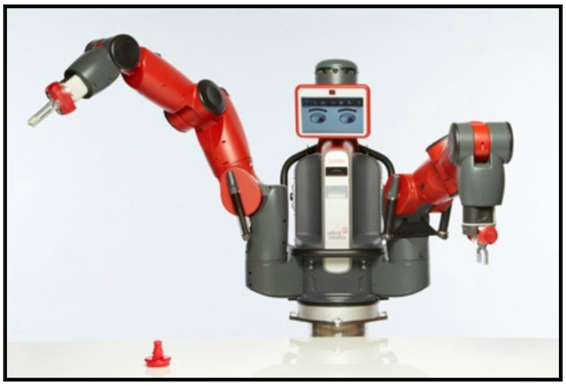
The approaches to HRC that are commercially available today are predominantly systems that are safe by design [35]. This means they are designed to be harmless to humans by being light weight and have limited movement speed. Some approaches propose some interesting strategies, however dominated by the limitations in the current safety standards [5], [6].

2.4.2 Commercially available approaches

Rethink Robotics™ have developed the Baxter system[1] (Figure 2.1a). Baxter is a double seven-axes arm, with a fully integrated control system. It can be installed in one hour and does not require any safety installations beyond the built-in safety system. But with only 2.3 kg payload per arm the work is limited to very light operations.

ABB's YuMi[2] is similar to Baxter, safe by design (Figure 2.1b). They represent a young generation of robotics, aimed at the collaborative market. However, as with Baxter, YuMi's payload is only 0.5 kg. This limitation inherently limits its possible uses. Further, even though

[1]Rethink Robotics™, www.rethinkrobotics.com (Accessed 11/10/2015)
[2]ABB, www.abb.com/robotics/yumi (Accessed 11/10/2015)


(a) Rethink Robotics' Baxter.

(b) ABB's YuMi.

Figure 2.1: The Baxter (a) and YuMi (b) lightweight collaborative robots are safe by design.

it is not strong enough to injure its human coworker, it not enough for a fruitful collaboration. On the other hand, both the Baxter and YuMi allow developers to focus on task execution rather than safety. Again, the tasks they can execute are limited due to low payload and reach. The system will allow a more automatic synchronous collaboration.

ABB have introduced the SafeMove system which is designed to bring the operator closer to the industrial robot [4]. SafeMove operate with zones in which the operator can move safely, and allows a more efficient use of the robot (Figure 2.2). The robot will automatically slow down as the operator approaches, and go to a full stop if the operator is too close.

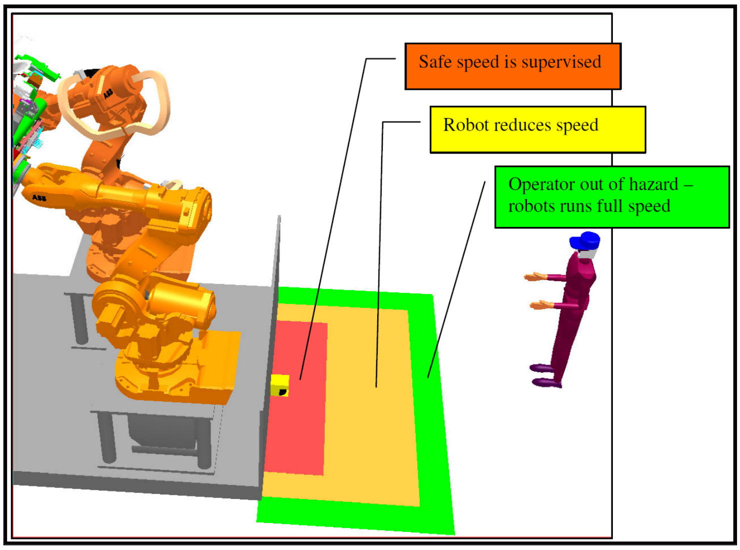


Figure 2.2: The ABB safe move system [4].

The SMErobots™ initiative has done extensive research on, and developed systems to simplify both the industrial level programming and safety issues related to industrial robot installations[3]. MRK Systeme have a commercially available safety system for selected low payload

[3]SMErobot ™, The european robot initiative for strengthening the comprehensiveness of SMEs in manufacturing, www.smerobot.org (Accessed 11/10/2015)

robots[4]. The system includes a capacitive cover for the robot, ensuring a full stop if the operator comes in contact with it.

A hand guiding approach to teaching have been commercially released by NACHI-Fujikoshi[5]. The technology behind it is not new, however, limited by safety standards, its been challenging for companies to commercialize technology from research. Now, with NACHI's lightweight robots which is safe by design, this approach can safely be presented commercially to the market.

2.4.3 Summary

There is no surprises in the commercially available systems. Although ABB's safemove and similar approaches allow some synchronous collaborations, the downtime for the human operator might be high, and the system is highly dependent on careful integration. Robots was initially designed and sought after for their strength and repeatability. The widely used safety by design approach is appropriate for many tasks that only require repeatability. However, for tasks that require the other benefit, strength, these approaches cannot be used. A system that works independently of the robotic hardware would be favorable.

2.5 Safety Related Approaches to HRC

2.5.1 Introduction

Safety related approaches are system that focuses on not injuring the human operator, a prerequisite for any HRC. As already presented, many approaches available in the industry today focus on safety by design. Therefore, this section will focus on approaches that can be applied to any robot. These approaches are often categorized as safety through control or safety by trajectory planning. However, in this research the strategy based on effect gives the categories and it is separated between reactive and proactive safety systems. It is not important for the human operator whether the robot continuously re-plans its trajectory or augments its pre-planned trajectory. It is thought to be of greater importance if the robot acts reactive or proactive against dangers. This is believed to have a greater impact on the effects of collaborating with a robot as presented in Section 2.2.

2.5.2 Reactive Approaches

Most contributions towards collision avoidance with redundant robots are based on static or kinetostatic images. Systems finding a collision-free joint space path [36], and problems that require maintaining end-effector constraints throughout the path [37]–[39] have been explored. It is usually distinguished between problems where a single goal is specified [40] and problems where the entire end-effector path is predetermined [41], [42].

[4]MRK-systeme GMBH, www.mrk-systeme.de (Accessed 11/10/2015)
[5]NACHI-Fujikoshi corp. www.nachi-fujikoshi.co.jp/eng/mz04/ (Accessed 11/10/2015)

The most fundamental strategy when it comes to safe robots is to stop the robot if a collision is detected. Approaches that measures the torque in the robots joints to detect collisions have also been presented [43]–[46]. This could also be an important safety barrier, however, these systems also interrupts the robots production and should only be regarded as a last barrier in a safety systems. Systems that avoid these collisions in the first place should be implemented in addition to these.

The second most fundamental strategy is to move away if too close to an obstacle. One of the first to present a system who's strategy was to maintain a distance to surrounding obstacles was Khatib in 1978 [47]–[49]. During the nearly 40 consecutive years there have been proposed numerous system based on the same strategy. The performance of the systems have naturally improved with modern computers and control theory. However, the fundamental strategy of moving away from an object that is too close, has not changed by much.

A virtual impedance control was presented by Tsuji et al. [50], [51]. In general impedance control, the robot moves based on external forces applied on the robot. The proposed method use virtual forces applied on the robotic arm based on information about the environment.

Balan and Bone [52] presented an efficient human collision avoidance system. The system used short term prediction models on both the robot and the human to reduce the effect of non-instantaneous response time. The systems strategy was is essence a separation monitoring system, ensuring a minimum distance between the human operator and the robot.

Kulić and Croft [53]–[56] have done extensive research on safe HRC and describe a safety system based on a danger index. The system uses the danger index in a real-time trajectory generator to re-plan its path if a danger threshold is exceeded.

Approaches with a high performance systems that avoid collisions by enabling evasive maneuvers if a danger is detected have been presented [57], [58]. Both systems avoided dangers, such as humans or objects, by moving away from them.

Lacevic et al. [59]–[62] presented an approach using Danger Fields based on kinetostatic information about the current situation. The Danger Field is a potential field spanned around the robots kinematic chain based on its velocity. The approach then used information about the humans position in the Danger Field to continuously minimize the danger in the current situation.

A concept using what was defined as a risk space was recently presented by Lo et al. [63]. A virtual impedance control was implemented in a reactive collision avoidance control for HRC. The system handled multiple possible collisions dynamically and proved to be very responsive. However, as will be discussed later, the design of the space is not truly based on the risk, rather a proximity or danger index.

Several more approaches based on the same strategy as these can be found [64]–[71]. Although their control strategy and thus performance vary and there is no doubt that there have been significant progress since the first proposed systems in the late seventies. Another important difference is what triggers avoidance behavior. While most detect the obstacles directly, some approaches calculations are more elaborate, using also the velocity of the robot or the human or both. However, they are all reactive obstacle avoidance strategies. The candidate have also presented a similar potential field approach as will be discussed later in this thesis [72], [73] (Figure 2.3).

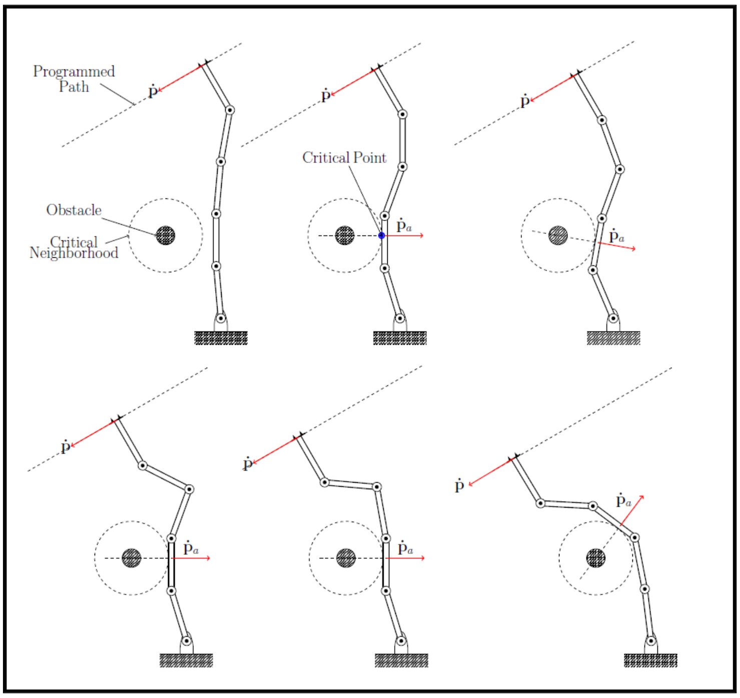


Figure 2.3: The basic strategy behind a majority of the reactive approaches [73].

2.5.3 Proactive Approaches

During the literature study in this research, no approaches than can be said to act proactively against dangers was found. Although, it is naturally to imagine that such a system would be beneficial. Several examples of the importance of trust and human attributes in HRC have been found, all of which promotes robots to act proactively against dangers. One can argue that some systems that re-plan the trajectory based on observations can be said to be proactive, however, the time span in which it is proactive is limited because they only address the current situation in their plan. It is not likely that the human would even notice this difference. Therefore, these approaches are not considered proactive in this research.

2.5.4 Summary

There is no doubt that the number of approaches to safe HRC with a strategy to maintain a minimum separation is plentiful in research. However, proper HRC have yet to be realized. This raises the question whether this strategy simply is enough. Is it enough to be merely safe? What is apparently lacking is a system that act proactively against the dangers, and not only reacts when it is detected that a danger is imminent. The importance of trust and human like traits in robotics presented in the literature in Section 2.2 supports the use of a proactive safety strategy as a new level of safety.

2.6 Task Related Approaches to HRC

2.6.1 Introduction

While discussing the safety of the collaboration, one must not forget the importance of the robot actually solving a task and serving a purpose. Many approaches focus on these task related challenges to realize HRC. As with the safety issue it is distinguished between reactive and proactive execution of actions related to tasks in this research. From the human operators perspective, this is an important aspect affecting the effect of working alongside the robot. The researchers in this category most commonly use robots that are safe by design and in that way avoid the safety issue.

2.6.2 Reactive Approaches

Haptic interactions in human-human collaborations was researched by Madan et al. [74]. The authors proposed five feature sets, including force, velocity and power related information. By discovering patterns in the haptic interactions, robots can interact accordingly with the humans in real time.

A system anticipating human activities, and selecting the correct response from the robot was presented by Koppula and Saxena [75]. The approach uses and anticipatory temporal conditional random field that models the rich spatial-temporal relations through object affordance as depicted in Figure 2.4). The most likely future scenario is calculated and an appropriate response form the robot is selected. A success rate at 85% was achieved.

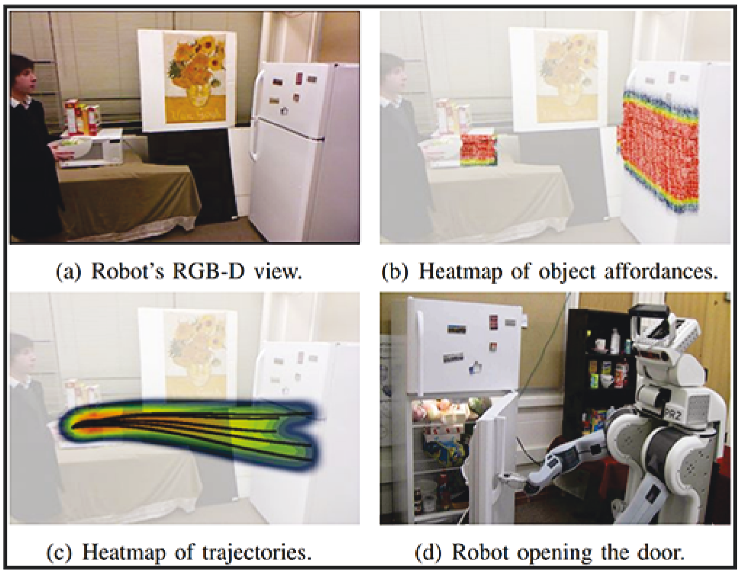


Figure 2.4: The Object affordances approach presented by Koppula and Saxena [75].

Billard et al. have presented systems for physical interaction between the human operator and the robot [76]–[79]. The systems are developed for teaching by guiding and in joint tasks between the human and the robot. This includes tasks where the human operator and the robot are lifting the same work piece. This is very interesting for operations that require both the strength of the robot, and the flexible fine position control of the human, such as assembly of heavy gearboxes. The approaches include the human in the control loop and use predictive models to improve the response of the robot. Other research projects have also presented similar approaches [80], [81]. Although the approaches control techniques and performance are different, the systems functionality are similar.

2.6.3 Proactive Approaches

Mainprice and Berenson [82] used workspace occupancy predictions based on an articulated motion library for the human. On the basis of the anticipated human task, the robot could plan its trajectory to solve its task. The approach relies on predefined human tasks and proper detection and selection of these tasks. The proposed system led to an efficient interaction between the human operator and the robot.

Kwon and Suh [83], [84] have presented a Bayesian network-based proactive human-robot interaction. The system use temporal and causal information to generate preparative actions to reduce the waiting time for the operator in a cooperative assembly task. The system demonstrated a significant reduction in task execution times, however, the approach were never concerned with safety.

A control framework proposed by Sisbot et al. [85]–[87] included a Human Aware Motion Planner for mobile robots. This component planed a robots path to equalize the cognitive load for the human and the robot. The system computed a point of transfer in a hand of scenario by evaluating safety visibility and human arm comfort.

The effects of working with a robot making anticipatory decisions in a HRC was studied by Hoffman and Breazeal [88]. The study demonstrated the importance of the robots behavior and reaction to the humans progression.

A study on human-aware motion planning demonstrated a system that planned the robots path on the basis of predictions about the humans next action [89]–[91]. The human's performance satisfaction and perceived safety were evaluated, and found to be increased using the proposed system.

Schrempf er al. [92] proposed a system that included uncertainties in the humans intentions when selecting an appropriate robot task in response. Two modules was used to realize the consept, an intention recognizer and a planner. When there was a high uncertainty in the humans intention, the system opted for proactive execution of tasks rather than idling.

2.6.4 Summary

Several well functioning approaches to task execution of different kinds exits in research. Both reactive and proactive systems are well represented in the literature. The systems are predominantly at a laboratory level, however showing very promising results. Although many of the approaches to proactive task executions may have a naturally positive effect on the safety issue when working properly, the system is not directly concerned with safety and would be unable to be affected by the safety situation if needed. Light weight robots that are safe by design are predominantly used in the presented approaches to task related approaches to HRC. Due to extensive research in this field, there is no direct connection to this in the problem statements. However, The robot should be able to maintain its productivity to as great extent as possible, while being safe.

2.7 Approaches Using Mobile Robots

2.7.1 Introduction

The focus in this research is on industrial robots, nevertheless several research projects investigate related challenges with mobile robots. Some of these approaches might produce insight and ideas adaptable for industrial robots.

2.7.2 Approaches using Mobile Robots

Meisner et al. [93] proposed an algorithm that utilized biofeedback from the operator to design human-friendly robot paths for mobile robots. The intention was to reduce stress in HRC by planning paths with a minimum distance to the operators. The strategy resembles the obstacle avoidance trajectory planning with industrial robots, however with biofeedback to measure the effect on the human.

An approach to path planning that computes the maximum velocity profile over a trajectory for a mobile robot have been presented [94]. The system use environmental information and robot dynamics to determine the profile. The result is a system where the mobile robot will e.g. slow down as it is passing a doorway, or plan its path further away from the doorway to keep a higher speed. This strategy is clearly adaptable to industrial robots in a system where the industrial robot plans its speed and path on the basis of where it is likely that the human operator will be.

A study aiming to give mobile robots the ability to find suitable locations for waiting was presented by Kitade et al. [95]. The system used a library of pedestrian trajectories and discovered suitable locations to wait not to disturb any humans. The ability detect areas with low human activity could be an interesting contribution in industrial robots as well to avoid disturbing the human operator.

A similar approach used predictive navigation by understanding human motion patterns [96]. By analyzing pedestrian trajectories, the system could select an appropriate action if a human was detected in its path. E.g., depending on where in the world you are located, people tend to go to either the left or the right when walking towards each other to avoid collision. A robot should have a similar behavior so that it can meet the humans expectations. Similar investigations could be conducted for industrial robots, and appropriate reactions could be selected if the robot have several options to avoid collision with the human.

2.7.3 Summary

Research on mobile robots have produced numerous interesting results. Even though the systems themselves might not be directly applicable to industrial robots, some of the basic ideas and strategies can be adapted. Path planning in close proximity of the human should include information also on which areas are frequently visited by the human, not only where a human is detected at this moment. The velocity throughout the path can also be manipulated to achieve safer operation in areas frequently visited by humans.

2.8 Problem Statements

As shown, there is a great deal of research going on in the HRC field. Many of which have produced very interesting results, however, there are many challenges yet to be solved. What is also clear is that no single strategy will realize proper HRC alone, a complex system with several modules would together make up an eventual HRC system. On the basis of this, a list of problem statements (PS) have been formulated in no particular order. The statements are criteria to be considered when developing a new approach to safe and productive HRC. The goal of this thesis is thus to fulfill the listed statements.

List of problem statements:

PS1: *The developed system should act proactive against dangers.* Today's safety systems moves the robot away if it is in a conflict with the human to avoid a collision. The developed system should avoid human-robot conflicts, thus acting as a new layer of safety.

PS2: *The developed system should be able to solve the necessary tasks to maintain its productivity.* The system should be designed to be independent of task and robotic hardware. Further, the developed system should have an awareness of what the human operator expects of it.

PS3: *The developed system should be designed to improve the effect the collaboration has on the human operator.* The developed system should reduce the workload for the human and have a low rate of false alarms. Further, it should inherit some human-like attributes to build trust.

The statements is a refinement of the problem formulation and will be used throughout this thesis. The first statement, **PS1**, is formulated on the basis of the literature review in Section 2.5 on safety related approaches to HRC. It was revealed that the the approach to safety in HRC have changed very little over the past decades, and a proactive approach was not to be found. As human being's behavior toward danger is proactive, we think before we act, it is expected that the robot also can benefit from this behavior. Secondly, today the safety strategies have two basic levels. The first level is to stop the robot if in contact with the human operator and the second level is to move the robot away if in conflict with a human to avoid contact. The proactive layer would be a third level who's goal is to avoid human-robot conflicts. Also related to Asimov's first law of robotics where it is stated that a robot cannot let a human come to harm through inaction. In other words, the robot should actively seek to keep the human safe, not just passively react if it is about to injure one.

The second PS concerns the productivity of the robot. While not injuring the human has the highest priority, if the productivity has no priority, it would be the safest to shut down the robot. Without productivity, the robot is meaningless. Therefore, while being a safety system, it should keep up the productivity to as great extend as possible. Therefore, as pointed out in Section 2.2, the robot should be aware of what the human expects of it. This awareness may allow it to solve the necessary tasks while being safe. Lastly, the proposed system should be designed independently of robotic hardware or task. The system should be designed in a way that that makes it applicable both for heavy-duty robots as well as robots that are safe by design. This way, resulting system may also be applicable to mobile robots, service robots and other branches of robotics research. The system should also be developed for a specific task or manufacturing process. Although the system might not be as efficient for all kinds of tasks, the developments should be carried out with this in mind.

Lastly, **PS3** states that the developed system should be designed to improve the effects the collaboration has on the human operator. This is closely related to the findings in Section 2.3 about trust in automation and human-robot teams. The importance of devoting attention to this aspect of the collaboration is clear as the human is as great a part of the collaboration of the robot. If the human operator is not comfortable with working with the robot, it is not beneficial for the collaboration. There is not one single solution to build trust in automation, however, it has been shown that the trust is closely related to interpersonal trust and several studies have shown a positive effect when implementing human-like attributes in the robot. In the development of the system, it should be investigated whether or not there are some human behavior that can be adapted to the new approach. In generality, the developed system should reduce the workload of the human operator. Lastly, the effect of a system that has a high rate of false alarms is clearly negative. The proactiveness should be exploited to investigate if it can have a positive effect on the rate of false alarms.

Moreover, in many other approaches the human's behavior is governed by strict guidance. The tasks are stylized and broken down into distinct components and are often overly articulated. There is a great uncertainty in how humans solve tasks. This pose a challenge in HRC which to some extent have been ignored in other task related approaches to HRC. This challenge can no longer be ignored and the system should in this research be tested without stylizing the human behavior. Therefore, the human operator should be given very open instructions in the experiments. The instructions should allow the human to choose how to solve the task, and even vary how the task is solved throughout operation. Realizing safe and productive HRC under these conditions poses an immense challenge for the robotic co-worker, while being an important barrier to breach.

Chapter 3

Situation Awareness and Risk Analysis in safe HRC

3.1 Introduction

It is evident that the safety of a HRC system is crucial, and that incorrect or inadequate operation has decidedly negative effects on the human operator's working capabilities. During the literature review, it was found that the system should have human-like attributes, in correspondence with **PS3** from Section 2.8. Analyzing today's reactive approaches to safety as a human attribute would simply be the reflex of pulling one's arm away from a collision. What is lacking is the decision not to move the arm into the area where a collision might occur in the first place. In other words, there is no decision involved in advance. If a robotic system is to be proactive against dangers and in correspondence with **PS1**, a decision making element should be implemented. Therefore, in this chapter the decision making mechanisms of the human will be discussed. The human being's risk perception is identified as the most important influence on the decision maker regarding safety. Current research and approaches to safe HRC will be discussed in light of the presented decision making model and the risk analysis framework. A novel approach to safe HRC is then presented and a model based on the decision making mechanisms in human beings is then proposed to realize the novel approach. How this model corresponds with the statements from the literature review summary in Section 2.8 is then briefly discussed.

3.2 Dynamic decision making in humans

The robot is in the literature more and more frequently expected to behave like a human to gain a HRC with equal trust and responsibility. Knowledge and literature about human-human teams are often applied to human-robot teams. Nikolaidis et al. [22] studied the effects of human-robot cross training, a human-human team tool where every member of the team learns the other members' tasks to better understand their own. The importance of trust in automation has been established and been revealed to be comparable to a human-human trust relationship [27]. One important factor in human-human relationships is the ability to predict each other's actions. People that behave unexpectedly, and often change their mind in the middle of an action may be hard to trust and be comfortable around. This has led to multiple studies to mimic human behavior in HRC, a strategy to gain the human operator's trust [20], [33], [97]. The reactive obstacle avoidance approach to safe HRC may also be described as a human attribute as we would also retract our arm or automatically move out of the way if necessary to avoid a collision with another human. What becomes apparent is that no single safety feature or task related feature will provide a truly autonomous HRC. Humans' ability to incorporate safety issues in their planning is vital in combination with the quick responses to sudden dangers.

Studying the available literature and current research on HRC, it becomes clear that safety strategies and task related strategies are almost exclusively separated. There are approaches that manipulates the robot with respect to safety, and others with respect to task related decisions. Lightweight robots are used in most HRC task related strategies, thus avoiding the safety challenge. Many safety strategies rely on reactive actions to avoid danger and are not concerned with task related challenges. Purely reactive safety systems will not alone provide the most optimal actions by the robot with respect to all aspects of the collaboration. The same basic strategy to safe HRC is applied in most research today, keep a minimum distance between the human and the robot. However, when studying the human's decision making mechanisms it becomes clear that just being safe at a given moment is not enough to be productive. The human's ability to make safe plans, both with regards to themselves but also related to others make them responsible decision makers, not just safe. A robotic system that shares these abilities to make responsible plans to keep the human from harms way should be further investigated.

One of the most fundamental attributes of a human being with respect to safety is the ability to judge risk [98]. Different people accept different levels of risk in their daily routines at work and at home. All day, in every situation, there exists a risk. Even in a simple activity such as crossing the road, the frequency and speed of the cars, the width of the road, and so on is observed. The person that want to cross the road use this data and compare them with his/her physical attributes, experience and how urgent or important it is to cross the road. The data goes through numerous psychological and cognitive processes and tells the person whether to cross the road or not. If the parameters tells the person to decide not cross the road, he/she might walk along the road to search for a change in some of the parameters. Along the road, authorities may have implemented measures such as traffic lights or zebra stripes to create safer crossings. This is all based on a risk analysis and the level of risk that the person is willing to accept.

Experience can tell us many things, maybe we are able to predict how many cars there will be on the road on the basis of the time of day. We know some roads are more busy during rush hour, and we go directly to a place where we know there is a zebra crossing, even though it might be detour. If we get to a road we identify as a highway, we might be more reluctant to cross, even though we cannot see any cars at the moment. The importance or urgency of crossing the road will affect the level of risk we are willing to accept. The accepted risk level might be very low if we are out strolling or walking the dog, we might not even consider crossing without a zebra crossing. If there is a slightly more urgent situation, such as being late for a meeting, we are willing to accept a slightly higher risk, and we start looking for a place to cross the road.

This risk perception enables the human to make decisions regarding safety, from a comprehensive understanding of the situation and predictions about the future status of the situation.

The mechanisms behind human's decision making with respect to safety is remarkable and is greatly influenced by the human's situation awareness.

3.2.1 Situation Awareness

In psychology, humans' comprehensive understanding of a situation is referred to as situation awareness (SA), and it is often discussed as a key element in dynamic decision making for humans. Many definitions of SA exist, however, that of Endsley [99] is frequently used. SA is there defined as: "the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future". Three levels of SA are then identified and placed in the context of decision making as shown in Figure 3.1.

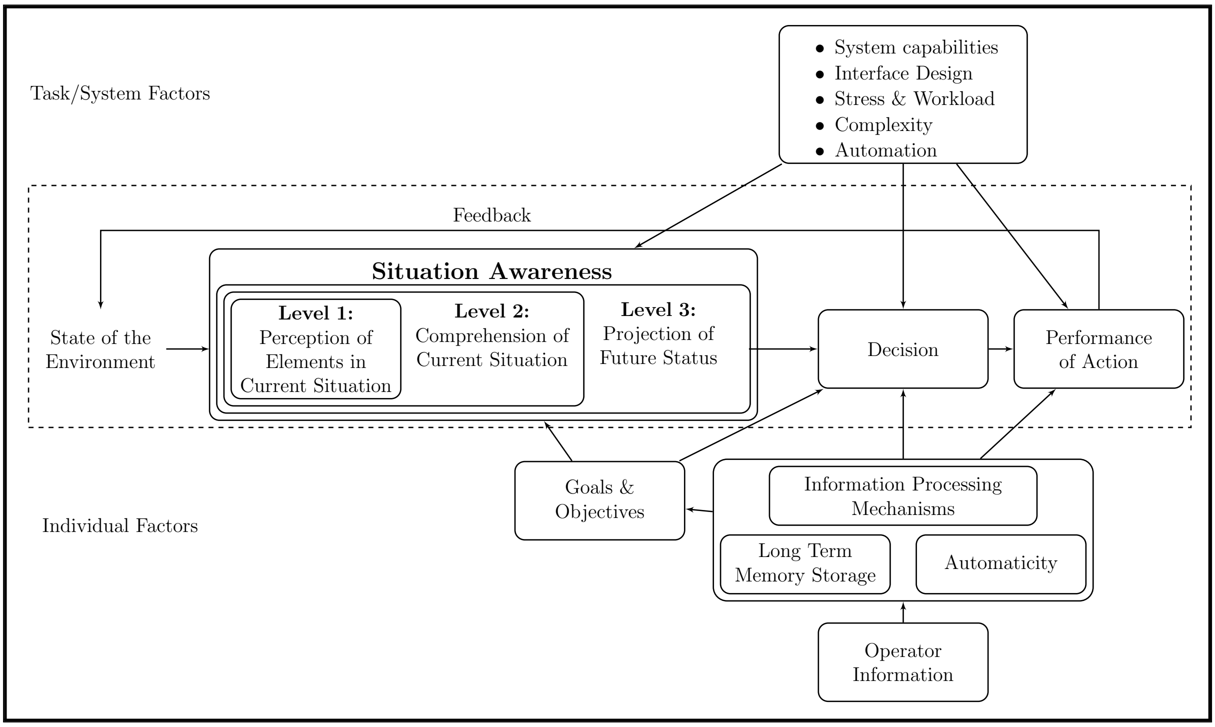


Figure 3.1: Model of situation awareness in dynamic decision making [99].

The model describes several factors affecting the decision making process. The main decision loop, including the SA element, is seen in the middle. Level 1 SA includes the perception of elements in a situation. An example with an air traffic controller will be used to better explain the different levels. An air traffic controller can with Level 1 SA observe multiple planes' current position and velocity. At Level 2 SA, new meaning is interpreted from these observations, such as the separation of the planes, and whether the planes are where they are expected to be at this time, etc. The highest level of SA, Level 3, requires some form of predictions about the status of the system in the future. Such as, where the planes will be at a given time in the future, will they collide at some point with their current trajectories. This perception, comprehension and projection highly influence our decision. A faulty outcome of the SA may result in an inappropriate decision later. Therefore, a high level of SA is vital in decision making. SA it-self is highly temporal and is not acquired instantaneously, but rather gained over time. One has to observe the changing situation to be able to reach Level 3 SA.

Further, an understanding about goals and objectives is also an essential element that influences the decision. The goals and objectives includes e.g. descriptions of a desired status of the system and are compared to the information from our SA. For the air traffic controller, the goals could be a minimum separation of planes or to avoid collisions. The purpose behind these two goals might be the same, to avoid planes from colliding, however, the level of SA will affect the decision and thus the effect of those goals. The effects of decisions made with Level 1 SA, Level 2 SA or Level 3 SA is shown in Figure 3.2. If the air traffic controller has only reached Level 2, a decision to take action and alter a plane's course, will only come when a separation that is too small is observed. This might result in a sudden change of course for one or both of the planes. If the goal was to avoid collision, simply comprehending the current situation might not be enough. This comprehension will only be able to see a collision that is already occurring. As a collision can be defined as a situation where the separation distance is 0, every separation above 0 will be regarded as acceptable and collisions become difficult to avoid. If the air traffic controller has a Level 3 SA, the outcome of the decisions might look different. The air traffic controller is now able to predict the status of the planes in the future, using their position, velocity vector, and knowledge about physics. If two planes are far apart now, but with their current course they will be within the minimum separation distance or collide in the future, the air traffic controller can earlier have them alter their course. Strictly speaking, both the cases with Level 2 SA and Level 3 SA are safe. However, it is apparent that the situation with Level 3 SA information is preferable. Being safe, is simply put, not enough, and the more responsible paths to the right in Figure 3.2 demonstrates for the pilots and passengers that the air traffic controller as the decision maker acts responsibly, and is trustworthy.

The event of two planes occupying space with too little separation is often referred to as an almost-accident, and must also be avoided. However, as illustrated with the air traffic controller, decisions made based on Level 3 SA result in smoother and safer flights. On the other hand, it is important to note that the Level 1 and Level 2 information is also vital. If for some reason the predictions at Level 3 are not correct, or the situation changes, the air traffic controller should take action if the two planes have too little separation.

Individual Factors will highly influence a human's ability to achieve a high level of SA. The human's attention level, experience and training lay a foundation for one's understanding and ability to perceive and comprehend the elements in a situation. Individual memory capabilities and cognitive functions highly affect the SA. A person who is forgetful will not be able to provide the same predictions of the future status as one that remembers well. Individuals have different skills when it comes to finding connections and comprehend different elements. Some might see a connection where others do not see any connection. All of these individual factors are highly influential in a person's ability to achieve SA.

Task/System Factors are factors that influence a human's ability to achieve SA. A complete list of these factors has yet to be determined, while some factors are apparent. The system's design and interface design are two factors which are relevant in a technological aspect. How a user interface is designed and how information is presented in that interface influences the

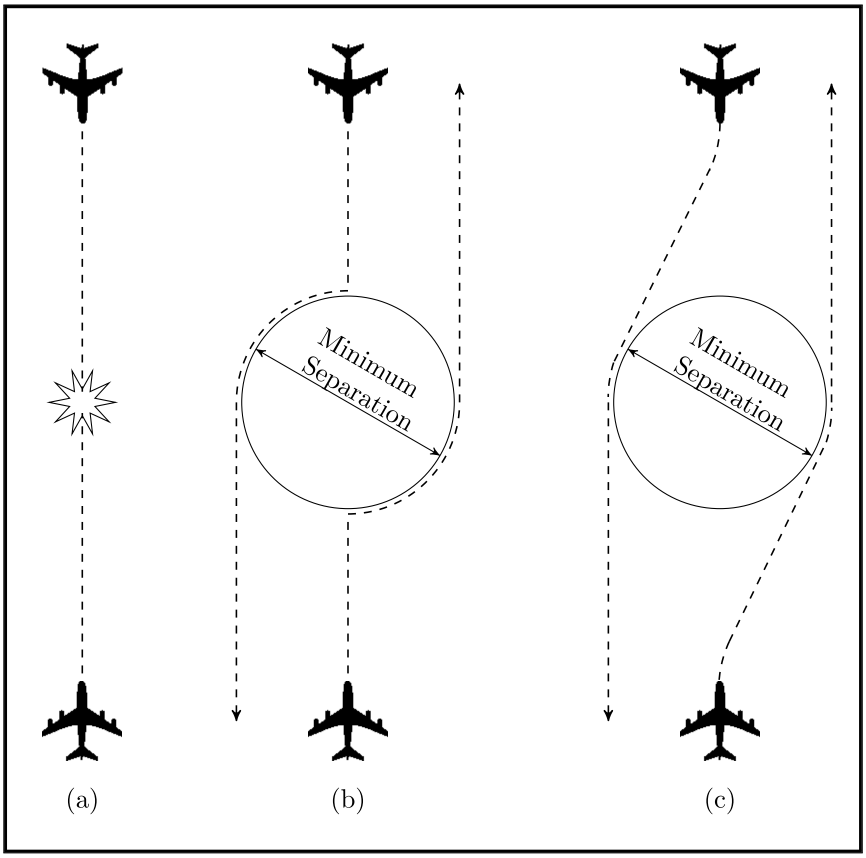


Figure 3.2: Resulting airplane trajectories with a goal of minimum separation based actions taken with with Level 1 SA (a) Level 2 SA (b) and Level 3 SA (c) information available.

human's ability to achieve a high level of SA. A complicated user interface that displays all available information at all times will not allow the operator to quickly gain a high level of SA, unlike a system that autonomously selects and presents only the necessary information based on the current situation.

3.2.2 Risk Analysis

As previously discussed, one of the fundamental attributes of human beings with respect to safety is risk perception. This has been modeled and adapted to industrial applications, and is now one of the most important tools when designing a safety system.

The goal of any safety strategy is essentially to reduce the risk. In this research, the definition of risk formulated by NORSOK [100]. In NORSOK Z-013N, risk is defined as a combination of the probability of an event and the consequence of the event. Further, a risk analysis is the process of using available information to identify possible accidents and estimate the risk. This definition is similar in most related standards. An overview of related components in safety and risk control is shown in Figure 3.3. A risk analysis is used to process the available information and estimate the risk via a consequence analysis and a likelihood analysis. The resulting risk picture is compared to a set of given risk criteria. The result of the evaluation determines whether or not more risk reducing measures (RRMs) must be implemented. The risk analysis is rerun with the updated system definition if more RRMs are implemented. The iterations stop and the system design is accepted when all risk acceptance criteria are met.

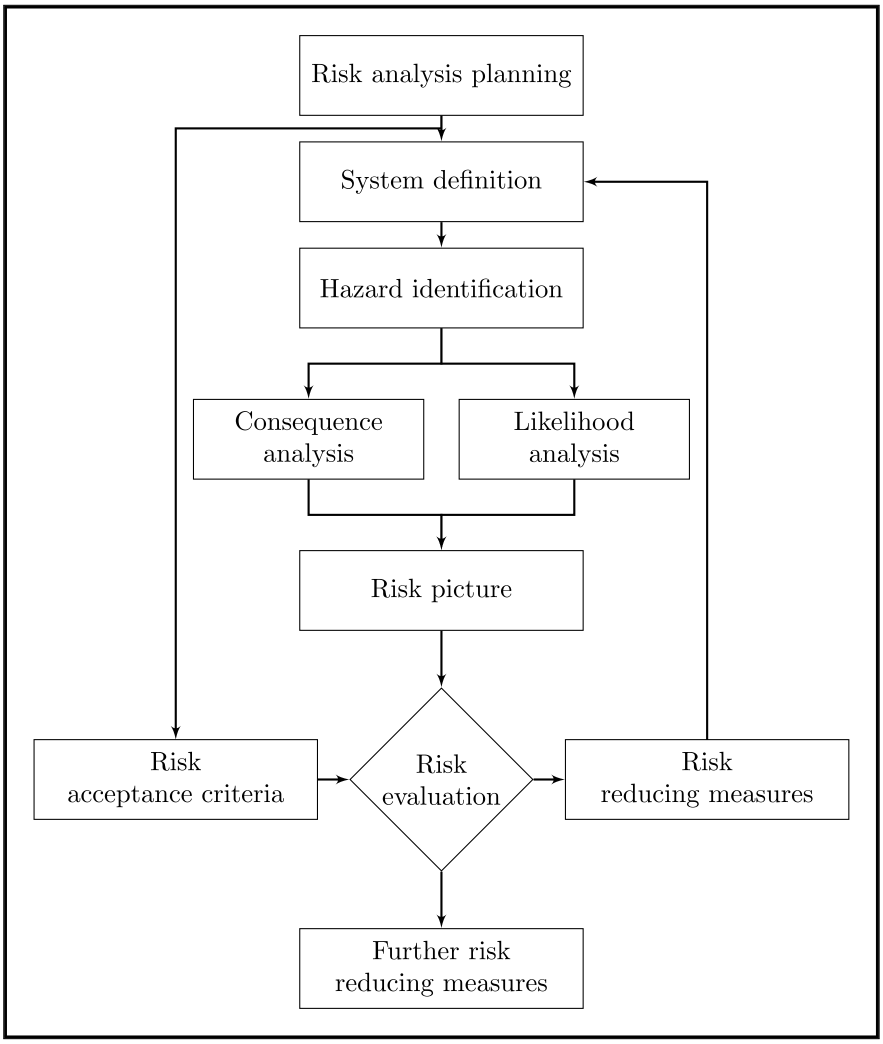


Figure 3.3: Risk estimates, assessment and analysis [100]

Understanding this framework is beneficial to understanding the mechanisms that affect situations with respect to safety. These mechanisms are important to understand when designing a proper safety system. The selection of RRMs are crucial to gain the desired effect and not to expose the system to other consequences. The most relevant components will now be discussed on the basis of the NORSOK Z-013N standard.

Hazard Identification

Identifying hazards is important to be able to analyzing risk. Hazards are unwanted events that result in consequences with a negative effect on the situation. Analyzing the road crossing example, the most apparent hazard is being hit by a car. Another could be tripping while crossing the road, or dropping what you are carrying.

Likelihood Analysis

The likelihood analysis involves estimating the probability that an unwanted event occurs. This will depend on several variables, and when involving humans, it is very challenging to properly estimate the likelihood.

When crossing the road there is a greater likelihood of being hit if the road is wide and busy, compared to a quiet narrow street. There might be a higher likelihood to trip if it is dark on a road full of potholes, compared to a smooth road at daytime. Whether you are stressed and are running, or calm and walking will also affect this likelihood. The type of shoes you are wearing, the weather conditions, your concentration level, and many other variables affect the likelihood of you tripping.

Consequence Analysis

The consequence is a complex entity itself. The consequences of an event can impact several aspects of a situation. It is often distinguished between personnel, environmental, and asset consequences. The personnel consequences include injuries, physical or psychological, or fatalities. Environmental consequences are negative changes in the environment from emission, waste, resource depletion and so forth. The consequences related to assets include material damage, loss of production, devaluation from bad reputation etc. The severity of a consequence differs greatly, and one is often forced to choose between consequences. Personal injuries and fatalities are for instance often considered more severe than material damage. Estimating the severity of each consequence is an important part of the analysis

The consequences are often very complex to analyze, and one consequence often alters the likelihood for a whole different set of events. If you trip while crossing the road, this would significantly increase the likelihood of being hit by a car. Similarly if you drop something in the middle of the road, and are forced to venture out in the road again. Being hit by a car opens a whole new set of possible consequences, a severe injury might affect your work and living situation. Therefore, consequences are difficult to quantify and a discrete set of consequence severities are often used in practical implementations.

Risk evaluation

When the risk picture is ready, it is possible to evaluate it by comparing it to the risk acceptance criteria. If the risk picture does not comply with the risk acceptance criteria, more RRMs must be implemented.

It is difficult to quantify both the risk and the acceptance criteria in the road crossing-example. The acceptance criteria will also differ greatly from person to person. Some might take their chance and attempt to cross a road with heavy traffic, while others don't cross even if they cannot see any cars. However, if the risk is within the acceptance criteria, the activity can be started.

What is apparent is that there is always a risk present, it might be ever so low, however, always there. The likelihood of an unwanted event may have been reduced significantly, while eliminating it is in most cases practically impossible.

Risk Reducing Measures

Risk reducing measures are the actions taken to reduce the risk. From the definition of risk, it can be said that an RRM either can be designed to reduce the likelihood of an event, or the consequence of that event.

The zebra stripes crossing the roads alert drivers that they must be aware of pedestrians attempting to cross the road. This will effectively reduce the likelihood of the pedestrian being hit by the car. Modern cars are designed to absorb energy in an impact and to avoid pedestrians from falling under the car if hit. Some researchers have presented pedestrian airbags and other pedestrian protection systems [101] [102]. These measures reduce the consequence of being hit, and the human might not suffer from as severe injuries. Some RRMs affects both the likelihood, and the consequence. A speed bumb will effectively reduce a car's speed, reducing the potential impact force, and increasing the time the driver and the pedestrian has to react. This measure effectively reduces both the likelihood and the consequence of the collision.

3.3 Risk Reducing Measures in Human-Robot Collaboration

As previously discussed, to reduce the risk, one can implement RRMs. In other words, a safety strategy involves a number of RRMs. In the context of HRC, four groups of RMMs can be identified (Table 3.1). These groups are either focused on the human or on the robot, and either on reducing the likelihood or the consequence of an accident. RRMs aimed at the human include fences, and light- and sound signals. These actions will remind the human of the danger associated with approaching a moving robot, thus making it less likely that the human would approach the robot. This is the most commonly used RRM today [3]. Another common approach is aimed at the robots. These include designing the robots to be less harmful, e.g., Rethink Robotics™'s Baxter[1] or the KUKA lightweight robot[2]. The system provided by MRK-systeme[3] ensures that if an accident occurs, it will have very little consequence from a safety perspective. Lightweight and slow robots will reduce the consequences of an accident. Another RRM is to implement control algorithms to automatically avoid a human, as in the research of Lacevic et al. [59], Petric et al. [58] and Kulic et al. [55]. These avoidance algorithms reduce the likelihood of an accident. The last RRM approach is probably never used in robotics. An RRM focused on the human to reduce the consequences of an accident would imply equipping the human with a helmet, armor, and other protective gear. Although a common approach in many other activities, it is extremely rare in robotics.

[1] Rethink Robotics™, www.rethinkrobotics.com (Accessed 11/10/2015)
[2] KUKA lightweight robot, www.kuka-labs.com (accessed 11/10/2015)
[3] MRK-systeme GMBH, www.mrk-systeme.de (Accessed 11/10/2015)

Table 3.1: Risk Reducing Measures for Robotic Systems

	Reduce Likelihood	Reduce Consequence
Human Focused	Fences and barriers Light and sound signals	Protective gear e.g. Helmet and armor
Robot Focused	Obstacle avoidance systems Proximity based emergency stop	Lightweight and slow robots Contact detection

A safety system will often be a combination of several RRMs. These are normally structured to be sequential based on which of the aforementioned consequences have the highest priority. With respect to only productivity and personal injury, it is obvious that preventing personal injuries has a higher priority. Robot-focused consequence reducing RRMs will effectively prevent personal injury but might cause an all new consequence: reduced productivity. This RRM should thus only be chosen when the type of production allows it.

More importantly, the RRMs can be seen in the light of the type of information it uses based on the levels of SA. As discussed in Section 3.2.1, the different levels of SA will influence the decision maker differently which will results in different types of actions. Level 1 SA provides information purely as perception of elements. RRMs in HRC that depend only on directly perceptible elements include emergency stop buttons, contact based emergency stop, door switches, light curtains and so forth as shown in Table 3.2. This equipment outputs a binary signal and the system use this signal directly to choose the appropriate action. However, it relies on a proper set of rules predefined by the human operator. E.g., the level 1 safety systems use simple rules "*if* E-stop=true *then* stop all motors". The consequence of this situation is that the production is stopped. Personal injury may be avoided, but on the other hand loss of production even for a few minutes is a dramatic consequence for most companies.

If the system is able to handle Level 2 SA information, new information can be derived from the perceived elements. The robot and the human operator's position can be combined to give a separation distance. A separation monitoring system can be implemented as an impedance control, using this separation distance information. A situation where a Level 2 SA action is called for now only augments the robot's task, or suspends it until the situation improves. Personal injury is avoided, however, the robot is operating close to the human, and the human operator's focus and concentration may have been compromised. The human operator is now alert and may be worried about the robot's next movement. This consequence is defined as a human-robot conflict, because the two are in conflict about whom should work in the given area at a given time. Similarly to with Level 1 SA, Level 2 SA is dependent on properly defined rules by the human. Rules such as "*if* human-robot separation<minimum separation *then* move away from human" is needed. A Level 2 SA safety system would not replace the Level 1 SA system, however be a layer of protection before Level 1 SA safety systems are needed. Without this level, one would have production stops far more frequent. The Level 2 SA thus reduces the number of emergency stops needed to provide a safe HRC system.

Enhancing the system's SA to level 3 opens up new possibilities to a safe and productive collaboration. With Level 3 SA comes the possibility to predict the future situation of the system. Combining this with the robot's tasks allows the system to plan its tasks from a safety perspective. The system becomes a proactive safety system because it can respond to a possible dangerous situation before it happens. Either by selecting tasks based on their safety profile, or by augmenting its path before it reaches a conflict with the human operator. As opposed to the importance of properly set rules for the Level 1 and Level 2 SA systems, Level 3 SA system could rely more on defined goals that are shared with the human operator such as productivity rates, and safety goals such as risk acceptance criteria. A perfectly functioning Level 3 SA system would not have any consequences, however, most situations are not perfect. Therefore, similarly to Level 2 SA, Level 3 SA is another layer of protection before Level 2 SA is needed. The goal of the Level 3 SA system is to reduce the number of human-robot conflicts. This would even further reduce the likelihood that an emergency stop is needed. Furthermore, reducing the number of human robot conflicts may allow the human operator to focus more on his/her task, rather than what the robot is doing.

Table 3.2 lists the desired priority of some RRMs. A proactive system can maintain the productivity to a greater extent than a reactive system. While an automatic emergency stop will reduce productivity, it is an important RRM to ensure safe HRC.

Table 3.2: Priority for Risk Reducing Measures

Level of SA	Risk Reducing Measure	Appropriate action	Consequence
Level 3	Proactive safety systems	Task selection, proactive planning	None
Level 2	Reactive avoidance systems, separation monitoring.	Augment task, task Suspension	Human-robot conflict
Level 1	Contact based emergency stop, emergency stop-button, door switch, light curtains	stop the robot	Production stop

In a system with all of these levels implemented the lower levels SA actions would overrule those of the higher level SA. If a task is started since the Level 3 SA system deemed it safe, it must still stop if it the emergency stop button is pushed.

The behavior of a system utilizing Level 3 SA information would much more resemble that of a human coworker. A human coworker would plan his/her actions based on your progression and share the responsibility for your team's safety with you. This approach would allow the decision making mechanisms in the robotic system to be influenced by the safety of the human, as opposed to obeying rules to protect the human's safety. This novel approach to safe HRC is introduced as Responsible Robots. This term has, to the candidate's knowledge, not been used previously.

The properties of the described types of RRMs are easily related to the SA levels in the human decision making model presented in section 3.2.1. By using the risk analysis framework and the human decision making model, a model that utilizes Level 3 SA in the robot's decision making process can now be designed to realize Responsible Robots.

3.4 Risk Analysis Based model for realizing Responsible Robots

Human's dependency on a high level of SA in decision making has been discussed. The risk perception has been identified as the tool human beings use to predict safety related issues in situations. Combining these to an approach to safe HRC by raising the system's SA to Level 3 by adapting the risk analysis model is therefore proposed. The novel model is an approach to realize Responsible Robots, introduced in Section 3.3.

The proposed model is shown in Figure 3.4. Some of the important components from the risk analysis model can be seen clearly, such as *hazard identification* and *risk estimate*. Other influential components such as *risk acceptance criteria* can be used directly as an objective in the *Goals & Objectives*-component. The *risk evaluation* will be included as a part of the *decision* component. The relevant components will be discussed further. The Risk Estimate component will enhance the systems SA to Level 3 due to the likelihood analysis as it holds the necessary predictions of the future status of the system.

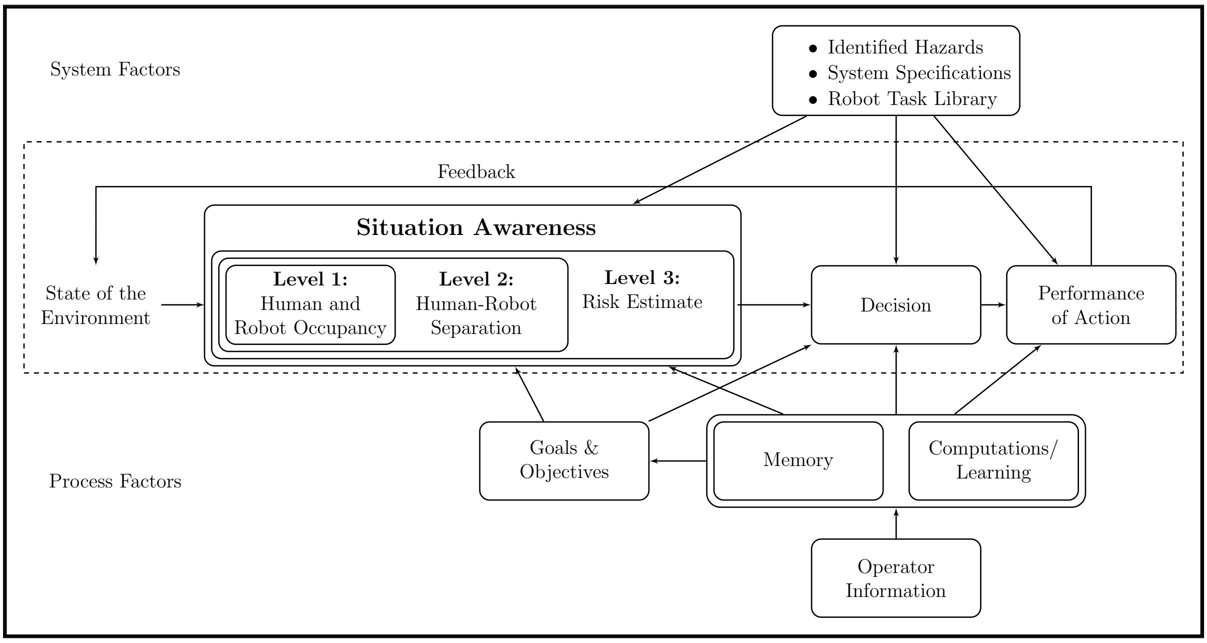


Figure 3.4: Situation Aware Risk Analysis Based Human Robot Collaboration Model

3.4.1 Decision Loop

The decision loop mainly consists of three components: SA, Decision, and Performance of Action. While many factors influence the decision, this loop illustrates the importance of SA. The risk estimate is now the active component at Level 3 SA as shown in Figure 3.4.

Situation Awareness

A safety system will depend on several RRMs which can be related to the different levels of SA as previously discussed. The systems at the different layers are all important to provide a safe and productive collaboration. The output of the SA component is thus composed of many types of information from all the levels of SA, all of which influence the decision making component.

Level 1: Emergency stop Information that is read directly from a sensory input is included at Level 1. This is mainly what is implemented in the industry today with sensors and input devices like emergency stop buttons, light curtains, door switches etc. Any input where no or very limited computation is needed makes up the safety system at Level 1 SA. This level of SA allows for very fast reaction time since the input doesn't need any computation. The importance of this level of SA must not be underestimated as this is usually the last barrier RRM in the safety system. The outcome of this level is very easily interpreted for the decision component such as *STOP=TRUE* or *STOP=FALSE*.

Level 2: Human-robot separation RRMs that require some comprehension of the situation is included at this level. Reactive approaches that augments the robots path based on the separation to the human is included here. The outcome of the SA at this level would be information such as the closest point on the robot to the human, the direction towards the human and the separation distance. The information needs some computation which puts more pressure on its design. However, most approaches to this today operate in real-time.

Level 3: Risk Estimate The risk estimate section of the risk analysis model will be implemented at this level of SA. A projection of the future status of the system is required, and the inclusion of a likelihood analysis in the risk analysis framework provides this. The risk estimation will consist of a likelihood analysis, a consequence analysis and a risk picture component. The risk picture provides information on the risk associated with performing a task, or occupying or moving in certain areas at this time or any time in the future. This data will aid the system in making safe decisions later. As this level is depending on predictions, more computations are needed than the other levels. Therefore, the provided information from this level will not update as frequently, however, its nature does not require such frequent updates.

Decision

The decision component represents the point in the system where an action is selected. This is influenced by all levels of SA, the system factors, and the process factors. The influences must be quantified and weighted in a computerized system.

Performance of Action

This component is used to evaluate the performance of the selected action. If the action of starting a task is selected, it would be natural to evaluate if the human and robot were in a conflict during the execution of the robot's task. If the system decides to augment the robot's path or configuration to maintain a minimum separation to the human, it would be natural to evaluate whether the minimum separation was maintained. The result of this evaluation can then be used to adjust parameters in the process factors. The risk threshold to start a task can be slightly increased if completing the task at that risk level caused a human-robot conflict. Parameters in the obstacle avoidance approach algorithms can be modified if the minimum separation is breached.

3.4.2 System Factors

The system factors are influences on the decision determined based on the surrounding system and task. This include variables and information that is set prior to operation, such as dynamic data for the robot, and positions of known obstacles in the environment.

Hazard Identification

The hazard identification component is corresponding directly to that of the risk analysis. Every event that will have a negative effect on the situation is included. In a isolated HRC case, every event that will harm the human is prioritized. This identification process is closely related to the tasks the human and the robot will collaborate, or work alongside, on.

System Specifications

Predetermined data about the robot such as the robot type, weight, stiffness, max speed, control bandwidth, and other data necessary to perform the computations and make the decision is provided here. This will also include the work cell's geometry, information about the production such as the geometry and weight of the parts and necessary tooling, and any other specifications that the system might need .

Robot Task Library

All the information needed to execute and make a decision about a task is made available here. What information is required to make a decision can be derived from the identified hazards and consequence analysis. This will include the robot's trajectory of each task, information about the sub-workspace, execution time, trajectory velocity and tool type and any other information needed for the robot to execute one of its tasks.

3.4.3 Process Factors

The process factor components in this model includes all the parts of the system that dynamically changes during operation.

Goals and Objectives

The goals of the robotic system is easily associated with Asimov's three laws of robotics. The first law says not to harm humans and can be more or less directly used as a goal. The second law states that the robot must obey the orders given by the humans. In other words, it must produce what is ordered by the human. A robotic system is meaningless without any productivity. Further, a prioritization is given by stating that the second law is not valid if in conflict with the first law.

With the risk analysis based approach to Level 3 SA, the most apparent objective relates to the *risk evaluation* in the risk analysis model in Figure 3.3. An important objective would thus be not to exceed a certain risk threshold. No action with a higher estimated risk than the threshold should be decided upon.

Objectives related to productivity would imply counters for the different robot tasks. This could also be related to observations of the human, the robot would be required to perform tasks according to the human's progression.

Memory

Every observation of the environment is stored in the memory. Unlike a human, a computerized system can keep detailed temporal and spatial data of the human's movements over a very long time span. Information about the task progression and other information needed by the other components are also stored by the memory component. The resulting parameters from the Computations/Learning component is also stored here and made available for the other components.

Computation/Learning

All analysis and interpretation of the observed data is processed in this component. How, and to what it is processed is highly dependent on the design of the other components. Computation of the currently observed data is required for the Level 2 SA comprehension, while the history of observed data available in memory is used in the computation required to reach Level 3 SA. The learning component will also execute the parameter adjustments on the basis of the evaluation from the performance of action-component. This component would mostly interact by reading and writing information to the memory component

Operator Information

If multiple operators are working with the robot, it might be useful to be able to distinguish them. Information that the system has acquired or might need about the operator is made available for the system in this component. At an enhanced stage, even bio-feedback from the operators may be stored, and the robot can adjust its behavior accordingly. A number of stress indicators might be used and the robot can adjust its behavior to reduce the stress of each individual human operator differently.

3.5 Discussion

The proposed model has many features that correspond well with the statements from Section 2.8. The robot will be enabled with an awareness of the risk related to its current pose, and planned path. The risk associated with executing the different tasks in the Robot Task Library is also possible to calculate. This opens up a series of new opportunities in HRC that is not depending on the type of robotic hardware (**PS2**).

Firstly, the system can be used in safe scheduling of the robot's task. Based on the risk perception, the system is able to determine when to execute the different tasks, and when to wait for the situation to improve based on risk acceptance criteria. In this way, the system is proactive in correspondence with **PS1**.

Secondly, in some cases the scheduler might not find a time to execute a task that satisfies the risk acceptance criteria. However, the productivity must be maintained, so at some point the task must be executed. If the situation is too risky to start the given task, the system can augment it by e.g. reducing the velocity of the robot. This will reduce the consequence of an impact, and the risk might be acceptable. If the risk is still too high it can make sound or light signals, give vibrotacktile feedback or in other ways communicate the risk to the operator [103]. Vibrotactile feedback is researched to help in teleoperations, and show a promising effect to deliver important information to the operator [104], [105]. This is in correspondence with **PS3**. Furthermore, instead of alerting the human about every single action the robot is taking, it can now decide when the human needs to be alerted. This allow the human to maintain focus, and be more alert when a signal is given. This will reduce the "cry wolf" effect of always giving signals about actions, and the human operator might respect the signal more. By being more selective of when an alarm is given, the system can correspond to **PS3**. Be reducing the total number of alarms given, it is natural to assume the number of false alarms will also be reduced.

Moreover, the estimated risk can be expressed as a vector field and used as a basis for a potential field or impedance control approach. If used in combination with redundant robots, the system could continuously search for the pose with the lowest possible risk [106]. This can be achieved in trajectory planning and real time velocity profiling. This approach would be a reactive response to proactive data, and can further enhance the safety of the human operator.

Furthermore, the goals and objectives component allows the system to have knowledge about what the human operator is expecting of it and how much flexibility is available to meet these expectations. This awareness of the human's expectations of it is in correspondence with **PS2**.

Lastly, the risk analysis could also help in analyzing the effects of other implemented RRMs [106]. The system could be run before and after implementation of RRMs or other design changes to the work cell to quantify the effect of them from a safety perspective.

Most importantly, by being proactive, the approach aims to reduce the number of human-robot conflicts by acquiring knowledge about them at an early stage. The research of Hoffman et al. [27] show how important trust in automation is, and how trust in machines can be related to interpersonal trust as stated in **PS3**. An important feature to maintain trust is to avoid unexpected behavior by the robot. Avoiding the situations where evasive maneuvers are necessary could help enhance this important trust since the model is developed on the basis of human's decision making process and risk perception and thus have human-like attributes.

3.6 Summary

The importance of situation awareness in the human's decision making process has been discussed in this chapter. Further, the three levels of situation awareness have been presented and related to current research and available safety systems. It was found that a proper safety system exploiting Level 3 SA was missing. Responsible Robots was introduced as a term to describe robotic systems with a Level 3 SA. These robots will make decisions that keep the human safe while being productive. Risk perception was identified as a means of enhancing the SA to Level 3 as the likelihood analysis give a projection of the future status of the system. The industrial standard risk framework was then presented and current safety systems was discussed in the light of the risk framework. A novel model to realize Responsible Robots was then presented. The model adapts the industrialized risk analysis framework into the human decision making model, thus enhancing the robotic systems SA to Level 3. The important components in this model was briefly introduced and discussed. The final section of the chapter then discussed how the proposed model corresponds to the thesis statements from Section 2.8 and some of the possibilities that comes with the implemented risk perception.

Chapter 4

Development of Model for Realizing Responsible Robots

4.1 Introduction

In the previous chapter, Responsible Robots was introduced as robots whose decisions are influenced by a concern for human safety. A novel model was introduced to realize Responsible Robots based on human decision making mechanisms as shown in Figure 4.1. In this chapter, the most important components in this model will be presented. This includes the Hazard Identification, Perception of Elements and Memory, Computations/Learning, Risk Estimate, Goals and Objectives, Robot Task Library and the Decision components. The likelihood analysis that is a part of the Risk Estimate component is the vital part that enhances the systems SA to Level 3. The other part of the Risk Estimate is the consequence analysis. While this part is equally important to realize an accurate risk estimate, it is far more researched than the likelihood analysis. While many researchers do not use the term "consequence analysis" directly, their results are close to directly applicable in the consequence analysis. This will be further discussed in Section 4.5.1. However, a likelihood analysis with this purpose is lacking in the literature, while also being the main component in realizing level 3 SA. Therefore, more attention has been directed at the likelihood analysis and its supporting components than the consequence analysis.

4.2 Identified Hazards

The Identified Hazards component holds information on all the possible hazards the system needs to be aware of. Although possible hazards in HRC are numerous, hazards causing personal injury is most often considered. Further, loss of production is an important situation to avoid. Other consequences like material damage are not considered part of the HRC problem in this research, rather general control and cell design. The reputation and other secondary consequences are also disregarded in the HRC safety problem in this research. These non-personal hazards may present a greater importance at a later stage in the development, however, hazards leading to personal injuries is the only focus of this research.

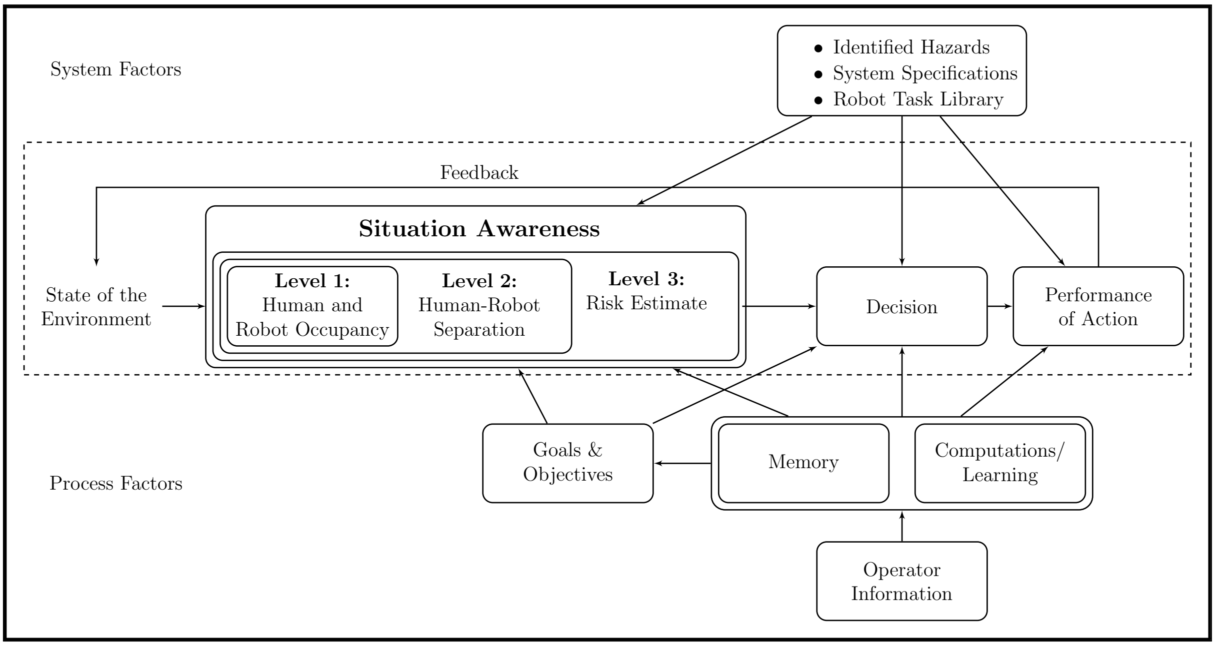


Figure 4.1: Situation Aware Risk Analysis Based Human Robot Collaboration Model for realizing Responsible Robots

The hazards related to human injury are numerous and depending on different types of collisions, clamping and stabbing, and depending on the tool. Haddadin et al. identified three types of impacts: unconstrained, partially constrained and constrained impact as shown in Figure 4.2 [107], [108]. Further, clamping and secondary impacts were also identified as possible hazards. A secondary impact occurs if the operator e.g. loses balance after an impact and falls. The fall will cause a secondary impact which might be more severe than the human-robot impact. Hazards related to operating with different types of sharp tools was also investigated [44].

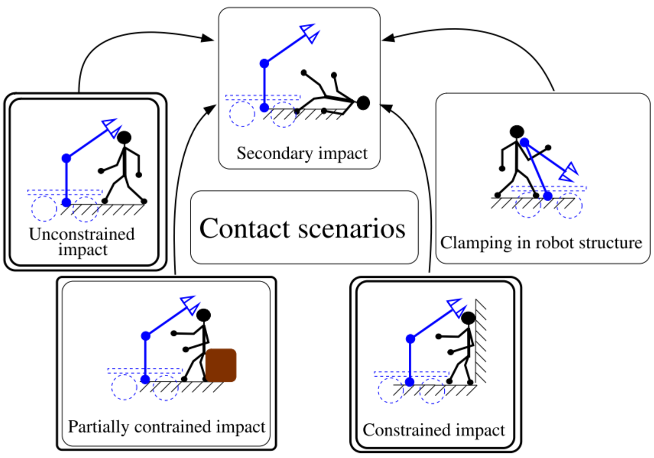


Figure 4.2: The five different contact scenarios [108].

What becomes clear is the vast number of possible hazards in a human robot collaboration. However, there is one obvious common denominator, the contact between the human and the robot. All hazards can be distilled down to a one simple hazard where the human and the robot attempts to occupy the same space at the same time.

Therefore, the input to the Identified Hazards-component is given by the human integrator prior to operation and the output are the unwanted events as seen in Figure 4.3. The design of other components will highly depend on these unwanted events. In reality these unwanted events are naturally embedded in the design of the SA-component.

Figure 4.3: The input and output of the Identified Hazards component.

4.3 Perception of Elements and Memory

The purpose of the Memory component is to serve as a storage unit for all the data needed by the other components. This includes, but not limited to, the temporal and spatial data of the human operator, and the resulting parameters of the Computations/Learning component.

The design of the perception of elements component is strongly linked to the identified hazards. Since the defined hazard is the human and robot attempting to occupying the same space at the same time, it can be derived that the position of the human and that of the robot must be observed. The position of the robot is found easily from the joint angles and the kinematic properties of the robot. The robot's position is not kept in the memory component, as this position is determined by the system itself. The human's position is on the other hand vital for the predictive capabilities of the system. The human's articulated position can be acquired by using a 3D depth sensor. Therefore, a depth camera is used to observe the human operator and a skeletal tracker provides several points along the skeletal structure of the human body. Each joint is represented as a point as depicted in Figure 4.4. If more points are needed to achieve the desired accuracy they can simply be interpolated from the observed points.

Now that the human operator's articulated position and its time stamp is known useful information can be computed. Since the identified hazard is the human occupying the same space as the robot, it is necessary to know the human's spatial and temporal position. Therefore, the workspace is divided into a cubic grid where the size of each voxel is determined by the required resolution of operation. Detailed work by hand on a workbench would require a smaller voxel size, while larger voxels would suffice in work where the human operator is moving around. The spatial and temporal data is also stored in the memory component. Furthermore, the data of interest are the times of when a human limb is visiting this specific voxel, Time of Visit (ToV). The time registered is the time when a human limb is entering the voxel. Even though it is occupied for a longer period of time, we just register the time the occupation starts (Figure 4.5). There is no need to register the entire occupancy as the system only needs to calculate the probability of a visit in unoccupied voxels. If a voxel is occupied, the robot can not work there.

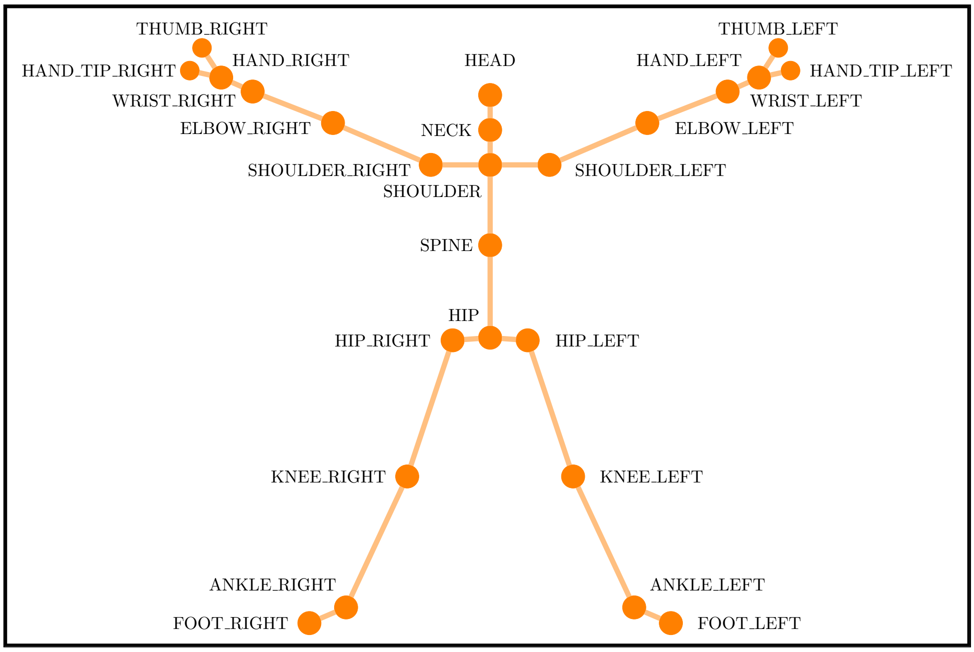


Figure 4.4: The tracked points along the human body provided by the kinect v2.

The ToV is observed in the absolute time space. The observations in the relative time space is the time since that last visit, Time Between Visits (TBV). The TBV is recorded as the time of departure from the last occupancy, until the start of the next occupancy as shown in Figure 4.5.

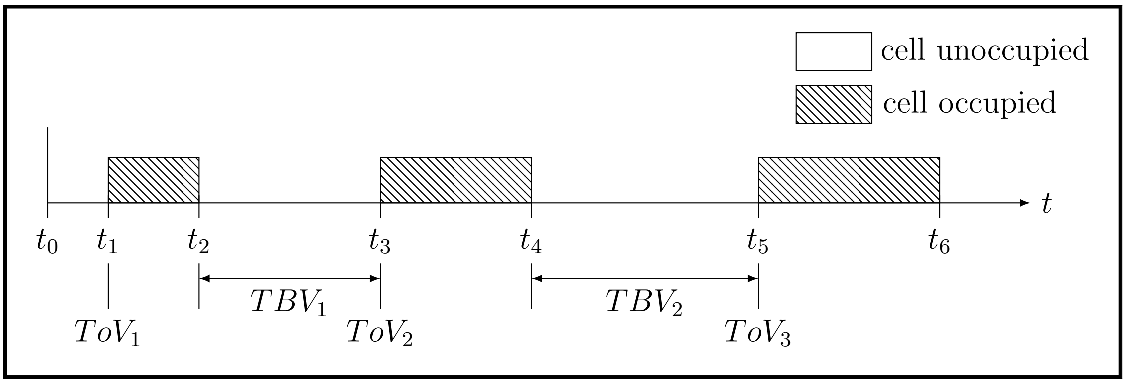


Figure 4.5: ToV and TBV measured from a voxel being occupied and unoccupied over time.

The dataset ToV[t] contains all the observations of ToV for one voxel, while TBV[t] contains the observations of TBV. For the observations in Figure 4.5, the data $ToV[t_1, t_3, t_5]$ and $TBV[t_3 - t_2, t_5 - t_4]$ can be extracted. Both of the datasets are organized in histograms; ToV(t) and TBV(t), where t is the time interval index. The datasets are stored in the memory and are thus available for the other components right away, however, the tracking continues, and the datasets in the memory are continuously updated so that the other components are always working with the most recent observations.

While there are several examples on human task recognition and classification [75], [82], a great part of the articulated human motion is not part of a task, or delicate finger work might be difficult to classify. Movement like scratching the head, checking the time, turning to see what that sound was, is generally ignored. This motion is most of the time unpredictable and inaccurate. Moreover, a tracked limb of the human will seldom take the exact same path during its next pass. Therefore, when a limb is tracked, the proximity factor κ is added to the appropriate bin in the histogram at the tracked point. At the tracked point, $\kappa = 1$. A reduced value of κ is added to the histograms of adjacent voxels up to a maximum distance, ρ, from the tracked point. This factor will also compensate for smaller errors in sensor data readings of limb positions. The resulting proximity factor, κ, is found by equation (4.1), where δ is the distance from the tracked point to the adjacent voxel. In effect, this turns the skeletal point tracker into a spherical representation of the human.

$$\kappa = \begin{cases} \frac{\rho-\delta}{\rho} & \text{if } \delta \leq \rho \\ 0 & \text{otherwise} \end{cases} \tag{4.1}$$

The memory also holds a history of the tasks executed by the robot to keep track of its productivity. To sum up, the inputs to the Memory component are the depth images with a skeletal tracking of the human operator and the selected action. The outputs are the TBV[t] and ToV[t] data sets, and the production history as depicted in Figure 4.6.

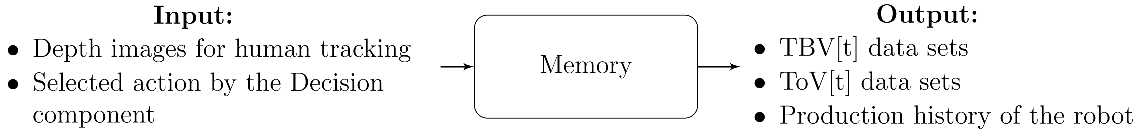


Figure 4.6: The inputs and outputs of the Memory component.

4.4 Computations/Learning

The Computations/Learning component interprets the observations stored in the memory. The component is responsible for all the calculations needed for the comprehensions for Level 2 SA and projections for Level 3 SA. The datasets available for the Computations/Learning components are the ToV[t] and the TBV[t] for each voxel. The information in the datasets must be interpreted to something the likelihood analysis can use. A regression algorithm is used to fit multipeak PDFs to each of the datasets. The parameters to the multipeak PDFs are then stored in the memory again, and is available for the likelihood analysis in the Risk Estimate component. The parameters are continuously improved, both with more observations and while there is no observable human. The inputs and outputs to the Computations/Learning component is depicted in Figure 4.7.

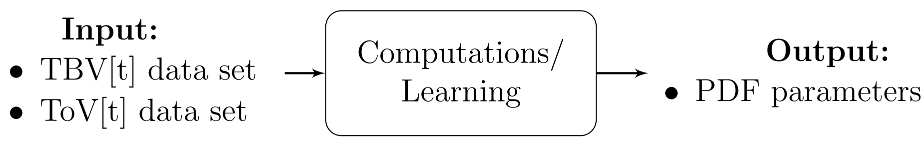


Figure 4.7: The inputs and outputs of the Computations/Learning Component

4.4.1 Regression Model

The regression algorithm will fit a multi-peak PDF to the ToV[t] and the TBV[t] datasets it retrieves from the Memory component.

As there is an unknown number of PDFs that is needed to get the best fit to the data, this can be said to be a variable size design space (VSDS) problem. The variable size lies in the number of PDFs. The dataset is first fitted to a single PDF, and the fitness is calculated as the mean square error (MSE). The algorithm is rerun, now with the available number of PDFs, g, increased by one, to two. The fitness of this double PDF curve is calculated and compared to that of one PDF. If the improvement in the fitness is lower than a set threshold, ε, the algorithm stops, and the parameters found with the single PDF is used. As long as there is a satisfying improvement in the fitness, the number of available PDFs are increased up to a maximum number of PDFs, G. The outline of the algorithm can be seen in Table 4.1.

For the non linear curve fitting itself, a well known Levenberg-Marquardt algorithm (LMA) is used [109]. The algorithm is a damped least squares method, which is considered to be slightly more robust regarding starting parameters, than other similar approaches. This is an important property, as the system can hardly make a qualified initial guess for every voxel.

Table 4.1: Curve Fitting

Step	Action
1	Set g=1
2	Run LMA, get $f_{i-1} = f_i$
3	Set g=2
4	Run LMA, get f_i
5	While $f_i - f_{i-1} > \varepsilon$ and $g < G$
5.1	Set $g = g + 1$
5.2	Set $f_{i-1} = f_i$
5.3	Run LMA, get f_i

Since the two datasets ToV[t] and TBV[t] have different properties, two different PDFs are used. A Gaussian distribution, $f(t;\mu,\sigma)$, is used for the ToV[t] given by (4.2)

$$f(t;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) \tag{4.2}$$

where μ is the location parameter, and σ is the shape parameter. And a Weibull distribution, $g(t;\lambda,k,\mu)$, is used for the TBV[t] which is given by (4.3)

$$g(t;\lambda,\sigma,\mu) = \begin{cases} \frac{\sigma}{\lambda}\left(\frac{t-\mu}{\lambda}\right)^{(\sigma-1)} \exp\left(\frac{t-\mu}{\lambda}\right)^{\sigma} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{4.3}$$

where λ is the scale parameter, σ is the shape parameter and μ is the location parameter.

The cumulative data and distributions are used since the datasets often are very chaotic and very difficult to fit a curve to. Let the Cumulative Distribution Functions (CDFs) of $g(t;\lambda,\sigma,\mu)$ be $g^*(t;\lambda,\sigma,\mu)$, and $f(t;\mu,\sigma)$ be $f^*(t;\mu,\sigma)$. Further, since multiple PDFs might be used to fit the curve, it is still important to ensure that the $CDF(t)|_{t=t_{max}} = 1$. A scaling multiplier, α, is therefore introduced to each of the distributions. The sum of all scaling multipliers for the active PDFs must be 1. The cumulative distributions are then given by

$$f^*(t;\mu,\sigma,\alpha) = \frac{\alpha}{2}\left[1+\text{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right] \tag{4.4}$$

$$g^*(t;\lambda,\sigma,\mu,\alpha) = \begin{cases} \alpha(1-\exp\left(\frac{t-\mu}{\lambda}\right)^{\sigma}) & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{4.5}$$

The combined distributions $G(t)$ and $F(t)$ are then given by (4.6) and (4.7)

$$F(t) = \sum_j f(t;\mu_j,\sigma_j,\alpha_j) \quad \text{where} \quad \sum_j \alpha_j = 1 \tag{4.6}$$

$$G(t) = \sum_i g(t;\lambda_i,\sigma_i,\mu_i,\alpha_i), \quad \text{where} \quad \sum_i \alpha_i = 1 \tag{4.7}$$

where i and j is the number of active PDFs in the two distributions respectively. The corresponding combined cumulative distributions follow naturally:

$$F^*(t) = \sum_j f^*(t;\mu_j,\sigma_j,\alpha_j) \quad \text{where} \quad \sum_j \alpha_j = 1 \tag{4.8}$$

$$G^*(t) = \sum_i g^*(t;\lambda_i,\sigma_i,\mu_i,\alpha_i), \quad \text{where} \quad \sum_i \alpha_i = 1 \tag{4.9}$$

To calculate the fitness, we use an objective function that gives us the Mean Square Error (MSE). The MSE for the two distributions are given by (4.10) and (4.11).

$$MSE_F = \overline{e_F(t)^2} = \frac{1}{t_{max}} \sum_{t=0}^{t_{max}} (F^*(t) - ToV(t))^2 \tag{4.10}$$

$$MSE_G = \overline{e_G(t)^2} = \frac{1}{t_{max}} \sum_{t=0}^{t_{max}} (G^*(t) - TBV(t))^2 \tag{4.11}$$

The system will continuously search for a solution that minimizes MSE_F and MSE_G. And while the tracking system will continue to add data to the memory component, and change ToV[t] and TBV[t], the refitting algorithm will adjust $F^*(t)$ and $G^*(t)$ to fit the new data. For some voxels, the fitted curve might not be a particularly good representation of the data. These voxels should be excluded in the later probability estimation, pending more observations or a better result from the regression algorithm.

4.4.2 Refitting

An important part of any forecaster is its ability to continuously learn, as more observations are made and the datasets in the memory component grows. There are in general two reasons to refit the MP-PDF to the dataset of a given voxel. The first being that new observations have been made in the given voxel, and the shape of the histogram have altered slightly. The second being that the curve fitted during the previous execution of the algorithm did not result in a satisfying fitness. While every voxel could benefit from being optimized again and again, it is very computationally heavy to run every voxel at every iteration. To get the best overall improvement of the system in a given iteration, the voxels go through a roulette wheel selection process. The roulette wheel selection is a tool where an individual's fitness is proportionate with its probability of being selected. In this case its reversed, the voxels with the lowest fitness will have the highest probability of being selected. This is done simply by negating the fitness of the fitted curve, as shown in equation (4.12). The selection is run until the desired number of voxels have been selected. This relearn rate should be carefully selected, not too low to avoid saturating the process with messy datasets, or too high, as the iterations will be very slow and many of the selected voxels might have very little room for improvement. A relearning rate of 25% was used in the later simulations.

$$p_i = \frac{-f_i}{\sum_{j=1}^{N}(-f_i)} \tag{4.12}$$

While voxels that have recently received new data might have a very good fitness before the new data arrive, they're given a priority in the selection process. The priority is given simply by manipulating their fitness for the sake of the selection. This priority is given, rather than simply recalculating the fitness, to give frequently visited voxels an advantage. As frequently visited voxels are in general more likely to be occupied by a human than an infrequent one, a good fitness in these areas is considered to be very valuable. This way of refitting also reduces the effects from the disadvantage of the required initial guess of the LM-algorithm since the previous solution is given as the initial guess.

4.5 Risk Estimate

The risk estimate is the means that will bring the systems SA to level 3. The risk is, as presented in Section 3.2.2, the multiple of the likelihood of an unwanted event and the consequence of that event. In all generality the risk can be expressed as the sum of these multiples as shown in (4.13), where p_i is the likelihood of event i and c_i is the consequence of that event.

$$r = \sum_i p_i c_i \tag{4.13}$$

As the likelihood is the part of the risk analysis that is responsible for the projections of the future status of the system, it will be devoted the most attention. Firstly, however, the consequence analysis will be presented in Section 4.5.1, followed by the likelihood analysis in Section 4.5.2. How the risk picture for the work cell is expressed is then described in Section 4.5.3.

The input and output of the Risk Estimate component is shown in Figure 4.8. The inputs are essentially the PDF parameters, data on the robot tasks and the objectives. The output is the risk associated with completing each of the robot's tasks.

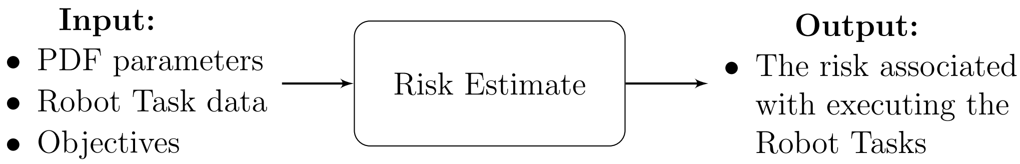


Figure 4.8: The inputs and outputs of the Risk Estimate Component

4.5.1 Consequence Analysis

Analyzing the consequence of an accident is a very challenging task that many have already researched. So many factors are in play, that it is close to impossible to calculate accurately. Simplifications are therefore necessary to gain reasonable results and many tools are available to generalize the severity of an accident. Marvel et al. have presented an approach to characterize hazards and their consequences in a HRC [13]. They also presented the data in Figure 4.9 which depicts injury criteria and body models from early drafts of ISO TS 15066 [7].

Body model Main and individual regions with codification			Maximum allowable Limit values of the injury severity criteria*		
Main body regions		Individual body regions	CSF [N]	IMF [N]	PSP [N/cm²]
Head with neck	1.1	Skull/Forehead	130	175	30
	1.2	Face	65	90	20
	1.3	Neck (sides/neck)	145	190	50
	1.4	Neck (front/larynx)	35	35	10
Trunk	2.1	Back/Shoulders	210	250	70
	2.2	Chest	140	210	45
	2.3	Belly	110	160	35
	2.4	Pelvis	180	250	75
	2.5	Buttocks	210	250	80
Upper extremities	3.1	Upper arm/Elbow joint	150	190	50
	3.2	Lower arm/Hand joint	160	220	50
	3.3	Hand/Finger	135	180	60
Lower extremities	4.1	Thigh/Knee	220	250	80
	4.2	Lower leg	140	170	45
	4.3	Feet/Toes/Joint	125	160	45

* CSF - Clamping/Squeezing force
IMF - Impact force
PSP - Pressure/Surface pressing

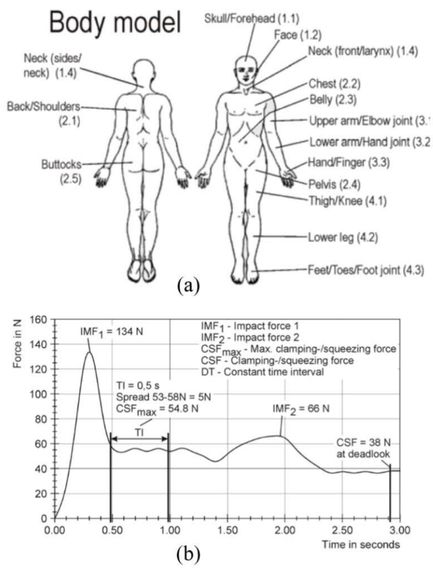


Figure 4.9: Injury criteria and body models from early drafts of ISO TS 15066 [7]. Clamping/Squeezing Force (CSF), Impact Force (IMF), and Pressing Cross Section (Pressure/Surface Pressing, PSP) limits are provided for several regions of the body (a), with the distinction between the two being characterized by duration and magnitude (b) [13].

Approaches that investigate and identify different contact scenarios in a Human-Robot related accident have been proposed by Haddadin et al. in [107] and [108]. They identify the five contact scenarios constrained, partially constrained and unconstrained impact, clamping and secondary impact. The last is often caused by one of the others, and might even be more severe. Some of these scenarios were tested in a crash-test during which contact forces, neck torques and other relevant data were collected. The data classifies the severity of the impacts. Haddadin et al. also investigated injuries caused by sharp tools on the robot, and developed a reactive avoidance system in [44].

Kulic et al. have presented a natural approach to measuring the danger [55]. They compute a danger index based on the potential impact force in a collision. Although they do not discuss their research as a consequence analysis, the consequence of an impact is closely related to the impact force. The velocity, stiffness and mass of the robotic manipulator combined with that of the human operator are the important factors in estimating the potential impact force.

There are also many standardized approaches to indicate the severity of an injury in non-robotic industry. The automobile industry started early to quantify injury severity and define indexes and criteria for analyzing the consequence of an impact [110]. Great parts of this research can be adapted to a consequence analysis. Most of this research is concerned with head injuries in automobile crashes. Some of the approaches attempt to relate the resulting head acceleration to the severity and likelihood of injury [111] [112]. The basis of these approaches is the Wayne State Tolerance Curve depicted in Figure 4.10. The acceleration of the head is related to severity of brain injury where points below the line are unlikely to cause brain injury.

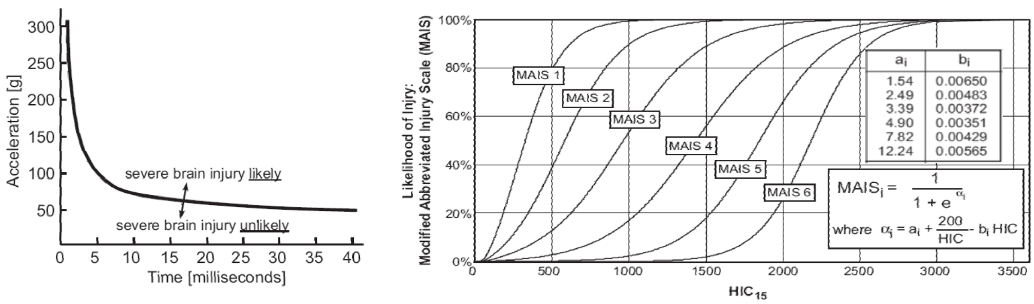


Figure 4.10: The Wayne State Toleance Curve: Points below the line are unlikely to be associated with severe brain injury; (right) Expanded Prasad/Mertz Curves: Chance of specific injury for a given HIC_{15} level [113].

Prasad and Mertz proposed a set of curves which statistically relates measured Head Injury Criteria (HIC) values to the severity and likelihood of a head injury (Figure 4.10) [113]. These curves can be combined with evaluated HIC values and then used to define the level of an injury resulting from a given head acceleration time history. The resulting injury indices can be used in a similar way to judge the severity of the injury combined with e.g. the Abbreviated Injury Scale (AIS). Using the AIS, any injury level is evaluated on a scale from 1 to 6, as shown in Table 4.2 [114].

It is apparent that a great deal of research and literature is available in relation to consequence analysis and that there are several aspects that can be included. However, the purpose of the consequence analysis remains the same: to quantify the severity of an unwanted event. The number of factors included in the analysis may vary based on the implementation and type of robot.

Therefore, in this model the consequence analysis will include the parameters Limb velocity, v_t, and Limb type factor, L, as seen in (4.14). A simplified model of critical areas of the human body is depicted in Figure 4.11, and is developed based upon the research of Haddadin et al. [107], [108] and Marvel et al. [13]. The torso and head are the most critical areas and the hands and feet are the least critical areas.

Table 4.2: Injury Severity Scale Classification according to AIS scale [114].

AIS	Injury level	Fatality Probability	Injuries Description
0	None	0%	Pain
1	Minor	0%	Light brain injuries with headache, vertigo, no loss of consciousness, light cervical injuries, whiplash, abrasion, contusion.
2	Moderate	0,1-0,4%	Concussion with or without skull fracture, less than 15 minutes unconsciousness, corneal tiny cracks, detachment of retina, face or nose fracture without shifting
3	Serious	0,8-2,1%	Concussion with or without skull fracture, more than 15 minutes unconsciousness without severe neurological damages, closed and shifted or impressed skull fracture without unconsciousness or other injury indications in skull, loss of vision, shifted and/or open face bone fracture with antral or orbital implications, cervical fracture without damage of spinal cord.
4	Severe	7,9-10,6%	Closed and shifted or impressed skull fracture with severe neurological injuries.
5	Critical	53,1-58,4%	Concussion with or without skull fracture with more than 12 hours unconsciousness with hemorrhage in skull and/or critical neurological indications
6	Maximum		Death, partly or fully damage of brainstem or upper part of cervical due to pressure or disruption, Fracture and/or wrench of upper part of cervical with injuries of spinal cord

$$c = L(1 + v_t^2) \tag{4.14}$$

4.5.2 Likelihood Analysis

The likelihood analysis is as previously emphasized an important piece of the model for realizing Responsible Robots. The result of the analysis is a prediction on the basis of the temporal and spatial observations stored in the Memory component. The Computations/Learning component have interpreted the observations and found appropriate parameters which will be used in this section. The probability of a visit in a voxel is assumed to be stochastically independent in the temporal domain, while the predictions in the spatial domain account for the probability due to a visit to an adjacent voxel. A prediction of the human's motion based on its recent motion in the spatial domain is presented first, then combined with the predictions in the temporal domain.

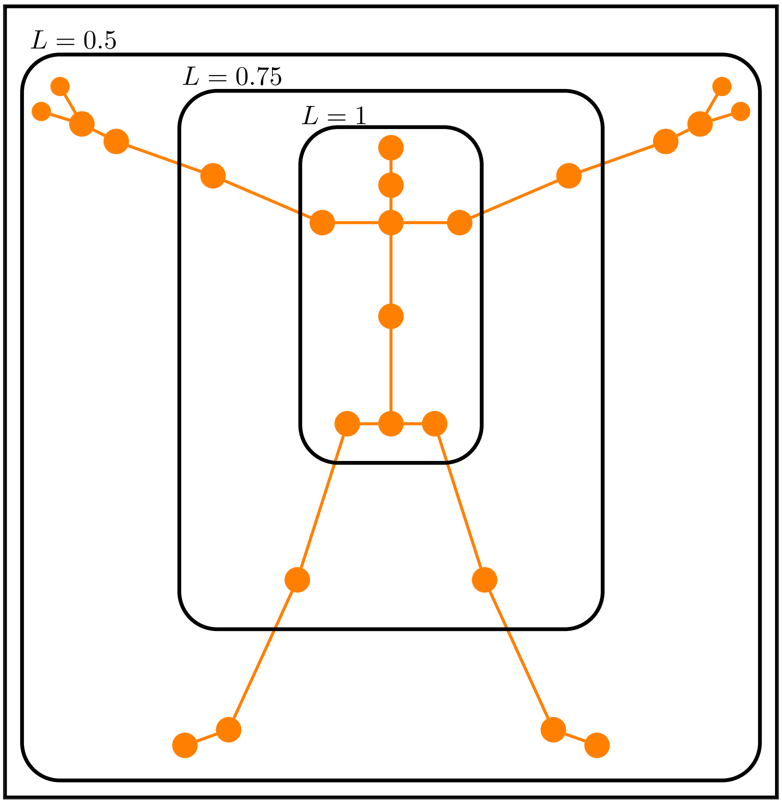


Figure 4.11: Simplified Limb Factor map with the head and torso as the areas with the potentially most severe consequence.

Human motion prediction

When calculating the portion of the probability triggered by the time of day or time since last visit, each voxel is considered to be stochastically independent. However, if a voxel is occupied, this will affect the probability that a neighboring voxel is or will soon be, occupied. A probability field is therefore constructed around the current observed position of the tracked points. The field is constructed around a predicted velocity vector based on a simple velocity prediction. The velocity prediction is a weighted average of the most recent tracking history. The prediction of the next iteration's velocity vector, $\mathbf{v}_t$, is found from (4.15). The number of iterations taken into account is denoted l, where η is given by (4.16).

$$\mathbf{v}_t = \frac{1}{\eta}\sum_{i=1}^{l}\frac{2l-i}{l}\mathbf{v}_{t-i} \tag{4.15}$$

$$\eta = \sum_{i=1}^{l}\frac{2l-i}{l} \tag{4.16}$$

The probability field is based on the Danger Field proposed by Lacevic et al.[61]. The author's approach to safe human robot collaboration included, what they called, a Danger Field (DF). This field was constructed to indicate the danger of occupying space in the vicinity of the robot. The danger can be interpreted as the likelihood that the robot would occupy a given voxel based on its current velocity. The field is a product of the distance from the object and the angle between the velocity vector and the vector from the traced point to the given point in space. It is given by (4.17), where $\mathbf{v}_t$ is the velocity vector predicted by (4.15). The object's observed point in space is given by $\mathbf{s}$, and $\mathbf{s}_t$ is an arbitrary point in space. The constants, k_1, k_2 and γ, are design parameters.

$$DF(\mathbf{s},\mathbf{s}_t,\mathbf{v}_t) = \frac{k_1}{||\mathbf{s}-\mathbf{s}_t||} + \frac{k_2||\mathbf{v}_t||[\gamma + cos\angle(\mathbf{s}-\mathbf{s}_t,\mathbf{v}_t)]}{||\mathbf{s}-\mathbf{s}_t||} \tag{4.17}$$

The field is expanded until $DF(\mathbf{s},\mathbf{s}_t,\mathbf{v}_t)$ is lower than a set threshold. A probability field is constructed by scaling the danger field, so that the sum of all active elements is 1 (4.18). PF_n is the probability that the tracked point, $\mathbf{s}$, will occupy point n, defined by $\mathbf{s}_t$, and corresponding to the danger in DF_n. This probability field gives the probability that a point in space will be occupied due to its adjacency to a currently occupied point.

$$PF_n = DF_n \left[\sum^{i} DF_i\right]^{-1} \tag{4.18}$$

Figure 4.12 shows the probability field and the relevant vectors. The field is rotational symmetric around the velocity vector.

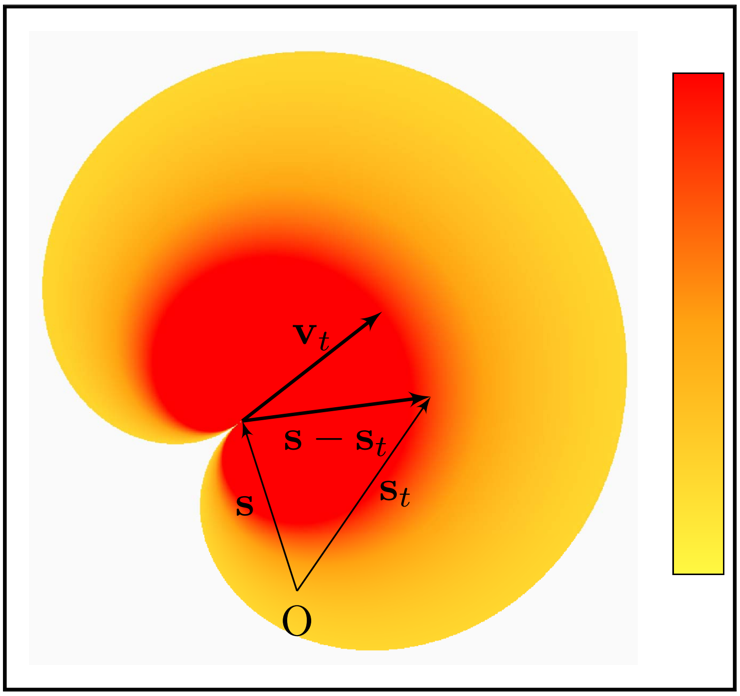


Figure 4.12: The probability field surrounding the velocity vector, $\mathbf{v}_t$. Red indicates a high probability of a visit, while yellow indicates a lower probability of visit.

Estimating the Probability

The predictions in the temporal domain are calculated on the basis of the multipeak PDFs, $F(t)$ and $G(t)$ whose parameters have been found by the Computations/Learning component. The parameters have been stored in the Memory component and can be used to calculate the probability that a human will occupy this space within a given time. Let t_V be the time of a visit, and Δt_V be the time between two visits. The probability of a visit $P(V|t)$ within a time, T, is then given by (4.19)

$$P(V|t) = P_{ToV}[t_0 \le t_V \le t_0 + T] \cup P_{TBV}[\Delta t_0 \le \Delta t_V \le \Delta t_0 + T] \cup PF \tag{4.19}$$

where t_0 is the current time, and Δt_0 is the current time since the last time the voxel was occupied by a human. PF is the probability caused by the prediction of the velocity of a human limb and its surrounding probability field. The probability of a visit due to the time of day, P_{ToV} is given by (4.20). Since there could be many visits per day to the same voxel, a binomial function is used to calculate the probability of more than one visit (4.20). The average number of visits per day is denoted ADV. $P_F[t]$ is found from the binomial equation as shown in (4.21). The probability of a visit due to adjacency, PF_n, is given by (4.18).

$$P_{ToV}[a \le t \le b] = 1 - (1 - P_F[t])^{ADV} \tag{4.20}$$

$$P_F[a \le t \le b] = \int_a^b F(t)dt \tag{4.21}$$

And the probability of a visit due to the time since the previous visit is given by (4.22).

$$P_{TBV}[a \le t \le b] = P_G[a \le t \le b] = \int_a^b G(t)dt \tag{4.22}$$

The integrals in (4.21) and (4.22) are solved easily using the CDFs from (4.8) and (4.9) as shown in (4.23) and (4.24).

$$P_F[a \le t \le b] = \int_a^b F(t)dt = F^*(b) - F^*(a) \tag{4.23}$$

$$P_G[a \le t \le b] = \int_a^b G(t)dt = G^*(b) - G^*(a) \tag{4.24}$$

This can now be expanded to cover more voxels. A region of interest can be defined based on the shared sub-workspace (SSW) required by the robot to complete a specific task. The SSW is available in the Robot Task Library Component as a part of the system factors. Let m be the number of voxels included in a SSW and T be the time the robot will stay within the SSW to complete its task. The probability of a visit, given we are in voxel n, at time t, is now, $P(V|t,T,n)$. The event of a visit in the different voxels can now be treated as stochastically independent because the PF term in (4.19) includes the probability of visit due to adjacency. The probability, $P(V|t,T,SSW)$, that a human will visit some part of the SSW within time T, given we are at time t, can therefore be found as shown in (4.25).

$$P(V|t,T,SSW) = 1 - \prod_{n=0}^{m} P(V|t,T,n) \tag{4.25}$$

The fitness of the fitted multipeak PDFs might be satisfactory for all the voxels inside a given SSW. This might be caused by too few observations to properly fit a curve, or that the regression algorithm simply could not find any proper parameters at all. These voxels are filtered out by setting a maximum accepted MSE_F and MSE_G denoted ε_{max}. If these voxels are rarely visited, the Computation/Learning component is likely to ignore it. However, if the voxel is frequently visited, it is more likely to be selected for a refitting as described in section 4.4.2.

4.5.3 Risk Picture

The risk picture is the output of the Level 3 SA and will highly influence the Decision component. The risk picture is both linked to the dedicated tasks for the robot, but also expressed as a gradient field as discussed in Section 3.5. The gradient field can be used in a potential field approach as part of an obstacle avoidance approach to reactive safety control.

Estimating the risk in a task

The risk associated with completing a robot task is the multiple of the estimated likelihood and the consequence. From the previous sections 4.5.1 and 4.5.2 the consequence C from (4.14) and the likelihood in (4.25) have been presented. The risk r_i associated with completing a specific task i from the Robot Task Library is expressed as the multiple of these as shown in (4.26).

$$r_i = P(V|t, T, SSW)C \tag{4.26}$$

The risk associated with completing every task in the Robot Task Library is calculated and passed on to the Decision component. The risk associated with the tasks are continuously updated on the basis of the latest information and observations available. This information is the key information when the system will decide whether to start one of the tasks, or wait until the situation improves.

Risk Field

The risk can also be derived as a risk field over the entire work space. Let T be a human limb whose position and velocity are defined by the vectors $\mathbf{s}_t = (x_t\ y_t\ z_t)^T$ and $\mathbf{v}_t = (v_{tx}\ v_{ty}\ v_{tz})^T$. The δ from (4.1) is then $\delta = \|\mathbf{s}-\mathbf{s}_t\|$, where $\mathbf{s} = (x\ y\ z)^T$ is an arbitrary point in space. . From the consequence analysis in Section 4.5.1 and likelihood analysis in Section 4.5.2, the risk related to a single point in space can be quantified. The risk field created by a single moving human limb at time i, $RF(\mathbf{s})_i$, is then defined by (4.27) with κ from (4.28).

$$RF(\mathbf{s})_i = \kappa L_f(1 + \|\mathbf{v}_t\|^2) + \gamma RF(\mathbf{s})_{i-1} \tag{4.27}$$

$$\kappa = \begin{cases} \dfrac{\rho - \|\mathbf{s}-\mathbf{s}_t\|}{\rho} & \text{if } \|\mathbf{s}-\mathbf{s}_t\| \leq \rho \\ 0 & \text{otherwise} \end{cases} \tag{4.28}$$

While multiple limbs may pose a risk at **s**, it is necessary to accumulate the risk posed by every limb. The risk field is thus expanded and derived by super positioning. $RF(\mathbf{s})_i$ is then the sum of the risk posed by all spheres for every limb (4.29). Each limb and sphere will have a specified ρ_t and $L_{f,t}$. Each sphere of the sphere-based geometry of the human is used to calculate the risk. The total number of spheres, l, and their radii, ρ_i, should be selected to best model the human body.

$$RF(\mathbf{s})_i = \sum_{t=1}^{l} \kappa_t L_{f,t}(1 + \|\mathbf{v}_t\|^2) + \gamma RF(\mathbf{s})_{i-1} \tag{4.29}$$

$$\kappa_t = \begin{cases} \dfrac{\rho - \|\mathbf{s} - \mathbf{s}_t\|}{\rho_t} & \text{if } \|\mathbf{s} - \mathbf{s}_t\| \leq \rho_t \\ 0 & \text{otherwise} \end{cases}$$

The field $RF(\mathbf{s})_i$ is by definition a scalar field. Nevertheless, a vector field can easily be constructed upon it using its gradient (4.30).

$$\overrightarrow{RF}(\mathbf{s})_i = RF(\mathbf{s})_i \frac{\nabla RF(\mathbf{s})_i}{\|\nabla RF(\mathbf{s})_i\|} \tag{4.30}$$

The $\overrightarrow{RF}(\mathbf{s})_i$ vector is anchored in **s** and with the direction of $\nabla RF(\mathbf{s})_i$. Its magnitude is set by the risk level in **s**, $RF(\mathbf{s})_i$. The risk associated with the robot's current pose can now be calculated on the basis of its occupied space, velocity vector and the Risk Field [106]. A reactive action can be taken if the risk is unacceptably high, such as alerting the human operator with an audio/visual output or other devices. The robot's path could also be augmented, very similar to a potential field approach. This approach, along with simulations, was presented in a paper by the candidate in a previous paper [106].

4.6 Goals and Objectives

The goals of the system is a summary of its purposes and is in this system closely inspired by the three laws of robotics, as mentioned in the introduction of this thesis. The first law states that "A robot may not injure a human being or, by inaction, allow a human being to come to harm". This is used quite directly in any safety system, as the goal in any safety system would be not to harm the human. Furthermore, "A robot must obey the orders given it by the human beings except where such orders would conflict with the first law". Therefore, for the robot to be useful and productive, some objectives regarding task progression and production rates must be included. However, the objectives regarding productivity must have a lower priority than the objectives to keep the human unharmed. The third law states that "A robot must protect its own existence as long as such protection does not conflict with the first and second law". From this there can be derived a goal about avoiding material damage to the robot or its surrounding. However, as opposed to the priorities in the three laws, this is given a higher priority than productivity, yet lower than the safety of the human being. Objectives to achieve these goals can then be formulated (Table 4.3).

Table 4.3: Objectives formulated for the Goals & Objectives component

Priority	Related goal	Objective
1	Human safety	The robot must stop immediately in case of contact with the human or an emergency stop etc is activated.
2	Human safety	The human and the robot must never have a smaller separation than the minimum separation.
3	Human safety	The system should never execute a task with a higher risk than the maximum accepted risk.
4	Material damage	The robot must stop immediately if it comes in contact with its surroundings.
5	Material damage	The robot must never have a smaller distance to its surroundings than a set minimum distance.
6	Productivity	The system must execute as many of the different tasks as is required within a given time span.

A simple objective to keep the human operator safe can be formulated on the basis of the information available from the Situation Awareness component. This component provides the risk associated with each of the robot's tasks as the Level 3 SA information. To keep the human safe, the system should never start a robot task that has a higher risk than a set threshold. An objective to at all times keep the risk below this threshold is thus implemented. From Level 2 SA, the human and robot separation distances is available. The second objective to keep the human out of harms way is thus to never go below a minimum separation distance. Lastly, the Level 1 SA information gives the opportunity to formulate a final safety objective. The robot must stop immediately in the case of an emergency stop, light curtain break etc. These three objectives, as summarized in Table 4.3, make up the objectives to keep the human operator safe.

The objectives to avoid material damage is not prioritized in this research. However, they much resembles the objectives to keep the humans safe. The only difference is that the human operator is replaced with the robot's surroundings, and prioritized after the human safety related objectives.

The objectives on productivity are very task dependent and will be described separately for each of the scenarios in the experiments. In some cases it might be required of the robot to perform a specific task based on the human operator's progression such as in an alternating assembly task. Other tasks might be more open, such as replenishing parts or other supporting tasks. However, an open objective is formulated for demonstration purposes.

Similar to the Identifying Hazards component the input of the Goals & Objectives component is given by the human operator/integrator prior to operation (Figure 4.13). The output is a list of the objectives in a prioritized order.

Figure 4.13: The inputs and outputs of the Goals & Objectives Component

4.7 Robot Task Library

The Robot Task Library component contains all the necessary information about the different robot tasks. This includes the robot programs themselves, the execution time and from this, the tasks sub workspaces. These sub workspaces are compared to the human operator's workspace, and the relevant SSWs are defined. The time from task initialization until the robot is clear of any SSW can be found from this, and stored in the Robot Task Library component. The robot's trajectories and velocities throughout the path will also be available here. In other words, all the information that is needed by other components about the robot's tasks. Also the Robot Task Library's input is added by the human integrator prior to operation and contains all the necessary data. The output is the same data in a format comprehensible for the system (Figure 4.14).

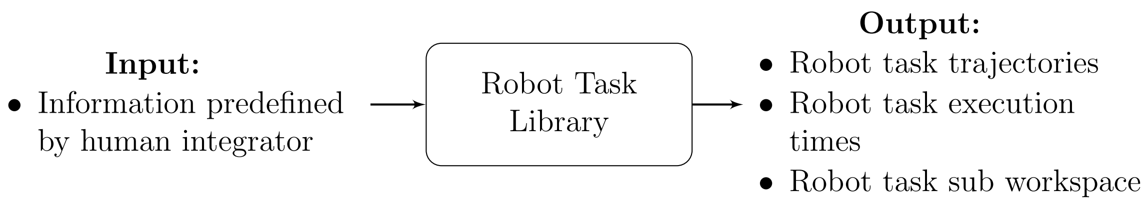


Figure 4.14: The inputs and outputs of the Robot Task Library Component

4.8 Decision

In all generality it can be said that the decision component combines all available information to achieve the goals and fulfill the objectives (Figure 4.15). The SA is as previously stated the most influential component. The decision component can best be described as shown in Algorithm 1. The highest priority goal is the goal to keep the human safe. The highest prioritized objective is based on the quickly accessible information from Level 1 SA, then the information from Level 2 SA. If those levels deem the situation to be safe enough to continue, and the robot is available, the system may execute a robot task according to the goals and objectives. In its simplest form the decision block goes through the objectives in a prioritized order and takes action to fulfill the highest priority objective currently not fulfilled. The decision components updates continuously, and a higher ranking objective might override an action taken to fulfill a lower prioritized objective if necessary.

Algorithm 1 Decision Component

```
procedure SELECT ACTION
    if E-stop = TRUE then
        STOP, Engage breaks
    else if Separation Distance < Minimum Distance then
        Augment Trajectory path and velocity
    else if Robot busy = FALSE then
        if Risk(Task_i) <Risk_max then
            if Task Penalty(Task_i) <Task Penalty_max then
                Start Task_i
                Task penalty(Task_i) ← Task Penalty(Task_i)+increment
            else Return
        else Return
    else Return
```

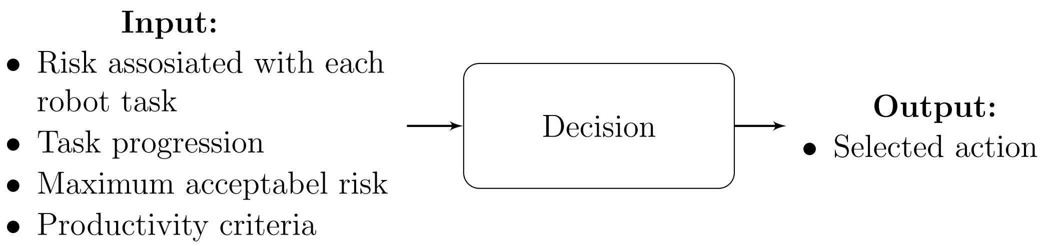


Figure 4.15: The inputs and outputs of the Decision Component

4.9 Other Components

The other components that are of relevance to the system include Performance of Action and Operator Information. The performance of action is an evaluation block that can be used to enhance the learning capabilities of the system. However, it is not implemented in the system at this stage. The focus has been on the implementation and influence of Level 3 SA. The operator information can also be used at a later stage to distinguish between the work style of different operators. In the later experiments, the system will be reinitialized for every operator. However, the system operator component could hold the necessary information to distinguish and automatically bring up the relevant observations from memory. The Robot Task Library has already been discussed, however, the remaining components in the system specifications contains more information about the system. This can include the cell layout with its known static obstacles and information about the production such as components' geometry etc. Any other information about the system that is needed by other components is available here.

4.10 Summary

In this chapter the relevant components in the Responsible Robot-model have been presented. The importance of the likelihood analysis have been emphasized as it is the component that accounts for the projection of the future status of the system, as required at Level 3 SA. Some of the components were identified as dependent on the specific collaborative work cell, and will be described in the experiments chapters. Furthermore, systems for realizing the Computation/learning, Risk Estimate, Hazard Identification, Goals and Objectives, and the Decision component have been proposed. The model and its components have been developed to fulfill PS defined in Section 2.8. The following chapter will verify the systems behavior with respect to the PS.

Chapter 5

Skill and Performance of the Likelihood Analysis

5.1 Introduction

The previous chapter presented the necessary components in the Responsible Robots-model, and the importance of a high level of SA in decision making has been discussed. Since the likelihood analysis is the component that accounts for projecting the future status of the system and enhancing the system's SA to Level 3, it is tested separately. In **PS1** it is stated that the system should act proactive against dangers. The system's ability to be proactive is directly related to the likelihood analysis. A verification of the performance of the likelihood analysis is important due to the importance of this component. In this chapter, the likelihood analysis, and its depending components, is experimentally tested as a binary forecaster and evaluated to confirm the fulfillment of **PS1**. The experimental setup will be presented first, followed by the results and an evaluation of the results.

5.2 Experiments

A series of experiments were conducted to measure the performance of the likelihood analysis. The Responsible Robots-model was reduced to only include the likelihood analysis and the components it is dependent on. The system was tested through three scenarios, each a different type of Lego assembly task. In each of the scenarios, the human had a set of instructions, like picking Lego bricks, assembling Lego figures or take Lego figure apart. The instructions had a low level of detail which allowed for a variety of different ways to complete the tasks in the instructions. The human operator could choose how to complete the tasks, and also vary how the tasks were completed throughout operation. Lego bricks in two sizes were used, later referenced to as large and small bricks, in a variety of colors. The robot also had a set of tasks, like replenish bricks or pick up Lego figures. Each of the robot's tasks had a defined SSW and a task execution time associated with it, all stored in the Robot Task Library. The human worked one "workday" in each scenario while being observed by the system, defined as day zero. After the system had learned sufficiently, the human worked another 10 "workdays". The system estimated the likelihood for the human to occupy the different SSWs within the robot's task execution times. The human occupancy is also observed and used to evaluate the system with a Decomposed Brier Score (DBS) [115]. Further the skill of the system is measured with the Brier Skill Score (BSS).

5.2.1 Experimental Setup

For the experimental evaluation, a workspace was built around a desk as depicted in Figure 5.1. A Kinect V2 sensor was used to track the human operator. The code was developed in LabVIEW, with the HARO3D™[1] VI library for integration with the Kinect. Two off-the-shelf desktop PCs were used, one of which was dedicated to handle the Computation/Learning component. The desk was located between the human and the robot, allowing the human and the robot to work on the desk from each side of the table. The desk used measured 180 cm by 40 cm. The workspace was separated into human sub-workspaces (HSWs) and robot sub-workspaces (RSWs) based on their tasks. The areas where a HSW and a RSW are overlapping are the shared sub-workspaces (SSWs) as shown in Figure 5.1. Note that HSWs, RSWs and thus also the SSWs might also overlap themselves. The HSWs were fixed for all scenarios and defined as three equal parts of the desk, HSW1 being the leftmost part of the desk, HSW2 the middle and HSW3 the rightmost for the human. The RSWs were defined from the robots tasks in the different scenarios.

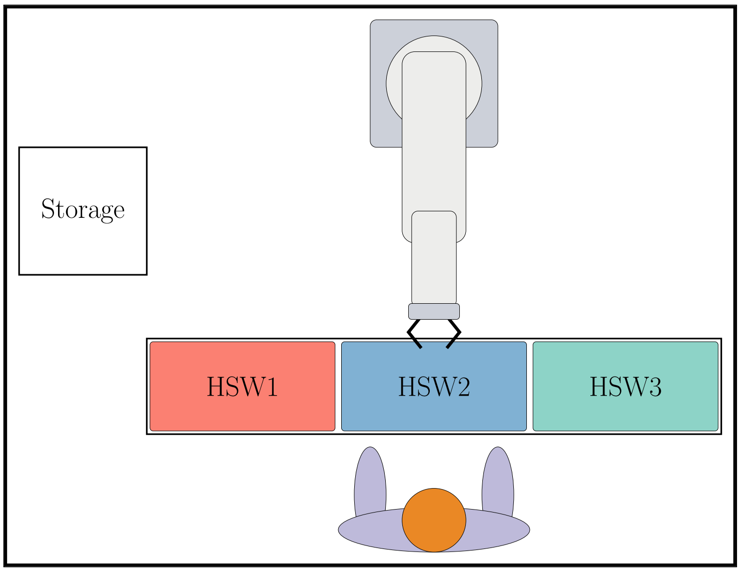


Figure 5.1: Layout of the workcell used in the experiments.

5.2.2 Procedure

The experiment was conducted with a human subject working in a total of three scenarios described in section 5.2.3. The participant first worked through one workday, designated day

[1] HaroTek LCC. www.harotek.com (Accessed 11/10/2015)

zero. The system was given some time for the computation/learning component to produce viable variables for the likelihood analysis. The participant then worked another day. This time, the system made predictions whether or not the participant would occupy a given SSW withing the corresponding task execution time for the robot. Since there were three SSWs in all scenarios, the system made three predictions each iteration. The system never acted upon the predictions thus continued to make predictions throughout the day. After the workday was completed, the DBS and BSS were calculated for that workday. Both during and in between workdays the Computations/Learning component keep improving the relevant parameters based on the new observations. The participant worked a total of 10 workdays in addition to day zero, in each scenario. After the 10 days in each scenario, the DBS and BSS were compared for each day, and its development analyzed. A summary of the steps in the experiment can be seen in table 5.1.

Table 5.1: Summary of the procedure the participant went through during the experiment.

Step	Procedure
1	Work one day as day zero
2	Allow system to compute
3	Work one work day while the system makes predictions about occupancy
4	Calculate DBS and BSS on the basis of the previous work day.
5	Repeat step 2-3 10 times
6	Repeat step 1-4 3 times, one for each scenario

5.2.3 Scenarios

The first scenario was a task where the human operator built several Lego cubes, and took them apart sequentially (Table 5.2). The robot's tasks involved replenishing bricks containers and picking up containers (Table 5.3) and resulted in three SSWs, two of which were overlapping (Figure 5.2). The robots task execution times were set based on an estimated time it would need from it decides to start the task, until it is clear of the SSW. The robot uses more time than the task time before it is ready to start a new task, however outside the SSW. In this scenario the human is visiting the different SSWs quite frequently and with approximately 180° phase shift. The amount of time the human is away from SSW1 does not leave much room for the robot to start its task. The human operator could typically complete 4 cycles during one workday.

In the second scenario, the human operator built a larger Lego figure, alternating between building in two different HSWs (Table 5.4). The robot's tasks involved replenishing bricks and picking up figures (Table 5.5) and resulting in three SSWs, all of them separated (Figure 5.3). The robots task execution times were set based on an estimated time it would need from it decides to start the task, until it is clear of the SSW. The robot uses more time than the task time before it is ready to start a new task. The human will be away from the two building zones for longer in this scenario, while more often and for shorter periods in the pick up zone. During one work day in the scenario 2 the operator could normally finish 3 cycles.

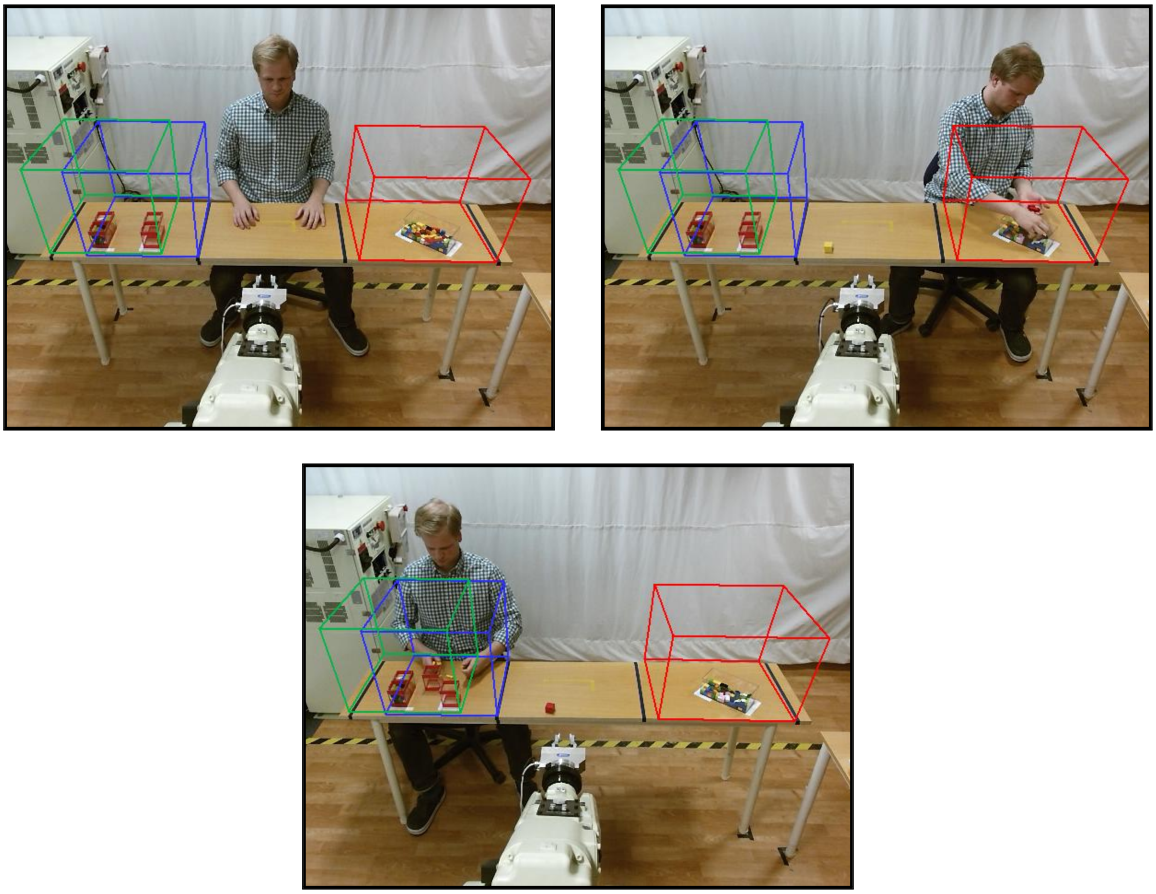


Figure 5.2: The workspace for scenario 1, with the three shared sub-workspaces indicated by the red, blue and green frames respectively.

Table 5.2: Human work instructions Scenario 1

Step	Action
1	Pick 3 large and 6 small bricks of the same color in HSW1
2	Assemble bricks to small cube in HSW1
3	Place the Lego cube in HSW2
4	Repeat step 1-3 a total of four times.
5	Pick up Lego cube from HSW2
6	Take Lego cube apart in HSW3
7	Place bricks in a container in HSW3
8	Repeat step 5-7 until no more cubes
9	Repeat step 1-8 until end of day

Table 5.3: Robot Tasks Scenario 1

Task	Action	Time
1	Replenish bricks in brick container in HSW1	25
2	Replenish empty containers in HSW3	24
3	Collect container with used bricks from HSW3	18

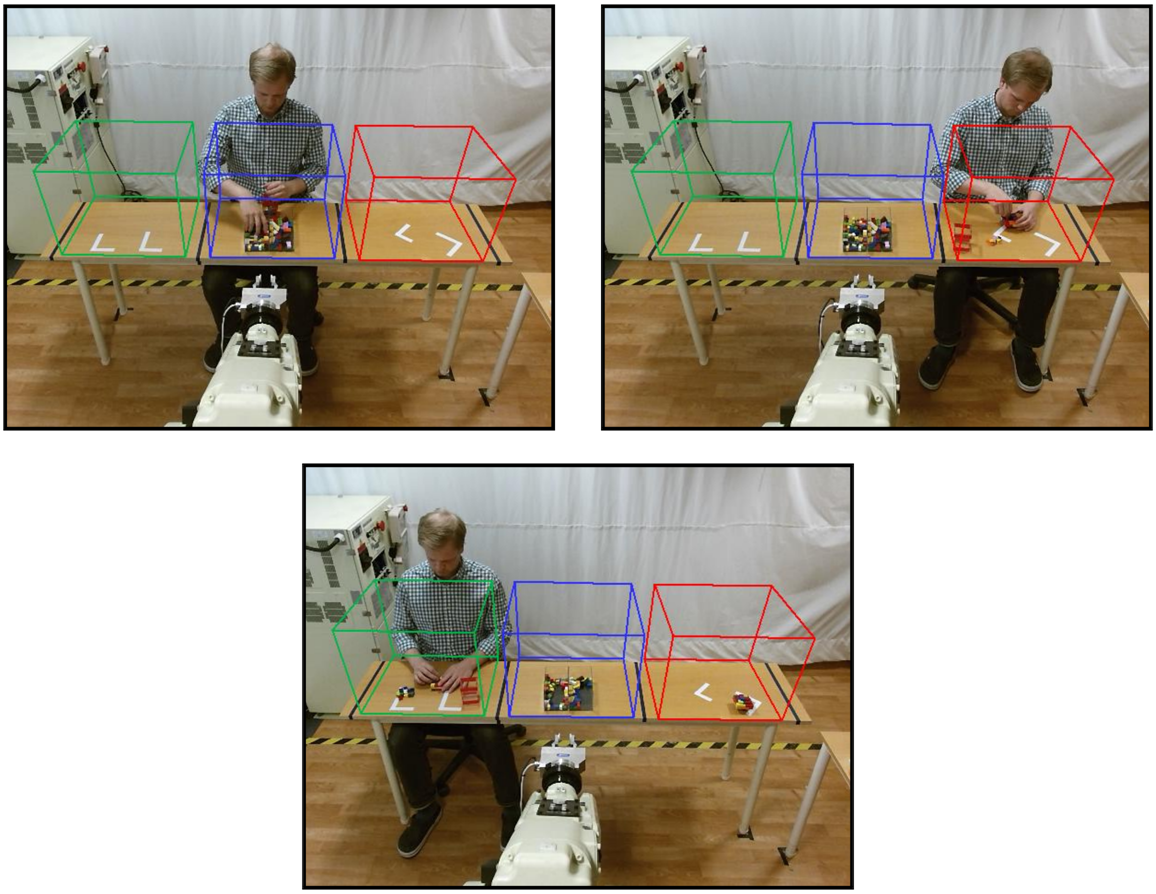


Figure 5.3: The workspace for scenario 2, with the three shared sub-workspaces indicated by red, blue and green respectively.

Table 5.4: Human work instructions Scenario 2

Step	Action
1	Pick 12 large and 24 small bricks of any color in HSW2
2	Pack bricks in an empty container
3	Bring container with bricks and assemble Lego figure in HSW1
4	Bring the now empty container to HSW2
5	Pick 12 large and 24 small bricks of any color in HSW2
6	Pack bricks in an empty container
7	Bring container with bricks and assemble Lego figure in HSW3
8	Bring the now empty container to HSW2
9	Repeat step 1-8 until end of day

Table 5.5: Robot Tasks Scenario 2

Task	Action	Time
1	Collect Lego figure in HSW1	16
2	Replenish brick in brick container in HSW2	25
3	Collect Lego figure from HSW3	16

The third scenario shows a human building two larger Lego figures at the time, then taking them a part. (Table 5.6). The robot's tasks were the same, and designated to the same three areas as in scenario 1 (Table 5.7 and Figure 5.4). The robot's task times were set based on an estimated time it would need from it decides to start the task, until it is clear of the SSW. Because of this, the robot uses more time than the task time before it is ready to start a new task. In this scenario the human would be building the figures in a sub workspace not shared with the robot, thus leaving all the SSWs unvisited at the same time. The cycles are also shorter, the operator usually finished 5 cycles in one workday.

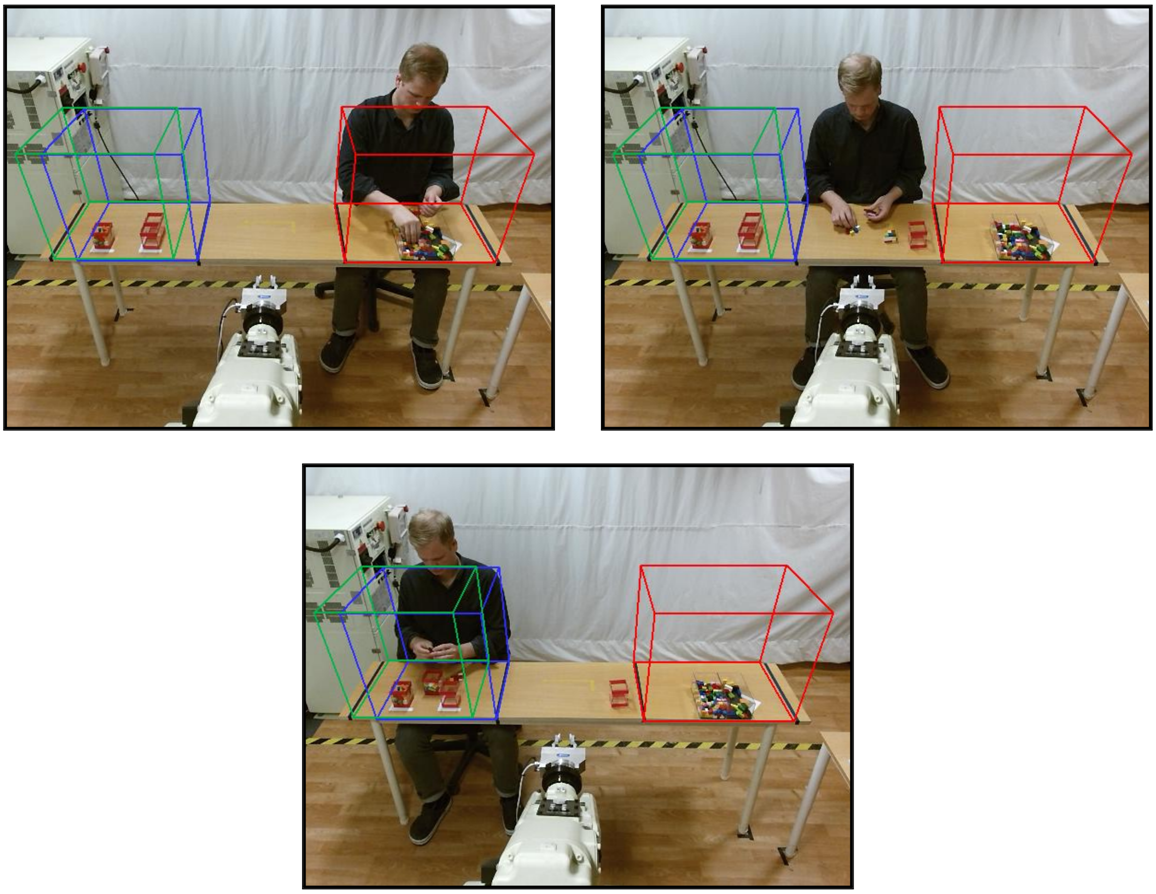


Figure 5.4: The shared workspace for scenario 3, with the three shared sub-workspaces indicated by red, blue and green respectively.

Table 5.6: Human work instructions Scenario 3

Step	Action
1	Pick 12 large and 24 small bricks of any color in HSW1
2	Pack bricks in an empty box
3	Bring the box with bricks and assemble two figures in HSW2
4	Bring the figures to HSW3 and take them apart
5	Pack used Lego bricks in an empty box
6	Repeat step 1-5 until end of day

Table 5.7: Robot Tasks Scenario 3

Task	Action	Time
1	Replenish bricks in brick container in HSW1	25
2	Replenish empty containers to HSW3	24
3	Collect container with used bricks from HSW3	18

5.2.4 Hypothesis

A set of hypotheses is formulated to determine whether or not **PS1** is fulfilled. Two hypotheses are formulated for these experiments, one regarding its performance and one regarding its improvement as the dataset grew. Firstly, it is hypothesized that the system is able to give a better prediction whether a given SSW will be occupied by a human within the given time or not than the base rate prediction. This means that the system has a low DBS and a BSS closer to 1 than 0. Secondly, that the system skill and score will improve throughout the 10 "workdays" the experiment lasts. This is measured by an decreasing DBS and an increasing BSS.

$\mathbf{H_{1.1}}$ The likelihood analysis is able to predict whether or not a human will occupy a given SSW within a given time.

$\mathbf{H_{1.2}}$ The systems performance is improving as more observations are made.

If both hypotheses holds, it is safe to conclude that the the system fulfills the proactive requirement formulated in **PS1**

5.2.5 Parameter Settings

The workday was set to 10 minutes in all scenarios, and the maximum TBV set to 1 minute. Both the ToV and the TBV datasets were distributed over 200 bins, giving a time interval of 3 seconds and 0.3 seconds in (4.10) and (4.11) respectively. The maximum accepted MSE for the fitness of the curve fitting was set to $\varepsilon_{max} = 5 \times 10^{-4}$. The maximum number of PDFs was set to 5 and the relearn rate was set to 25%.

The design parameters used in the human motion predictions, k_1, k_2 and γ, were set to $k_1 = 1$, $k_2 = 0.25$ and $\gamma = 1$. The length of the human velocity prediction history, L, was set to $L = 1second$. The maximum distance of the proximity field ρ is set to $\rho = 20cm$. The likelihood analysis iteration time was set to 1 second.

5.3 Results

The system could typically deliver a well enough learned system to continue 10-15 minutes after the day was completed. The scenarios typically generated enough observations to provide usable parameters in 1000 voxels. While a larger number of active voxels will result in a higher computational cost, the cost at this point is considered to be well within what is reasonable to continue with this approach.

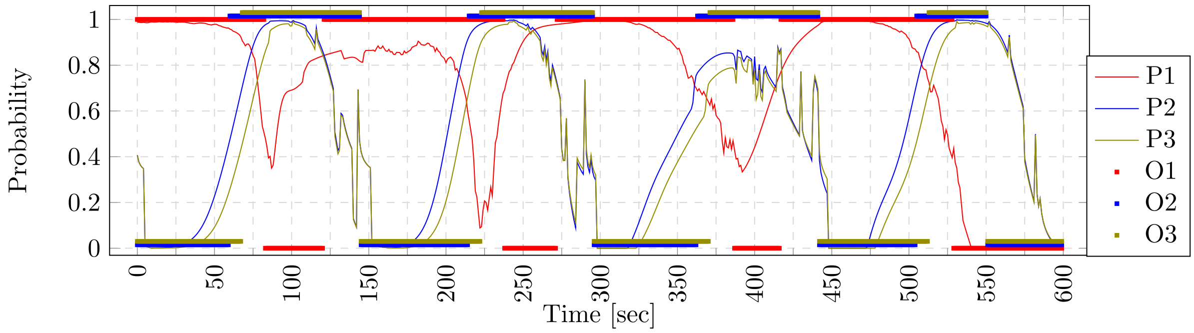


Figure 5.5: The predictions, P_i, through a day and the corresponding observations, O_i, for Scenario 1

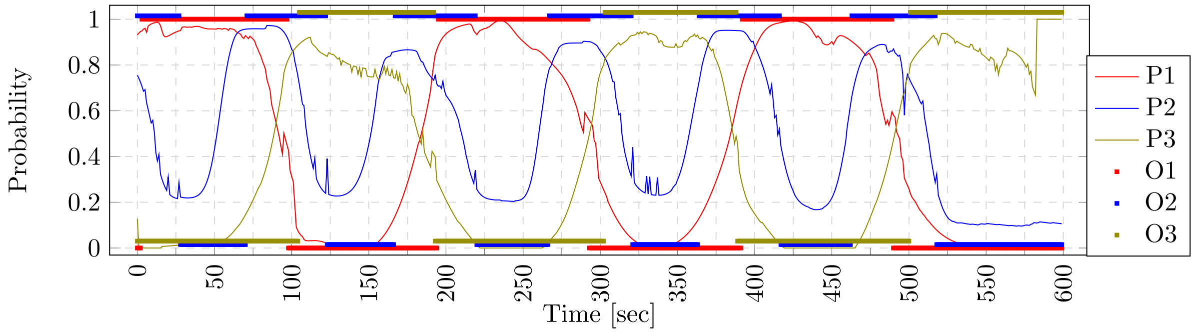


Figure 5.6: The predictions, P_i, through a day and the corresponding observations, O_i, for Scenario 2

The plots in Figures 5.5-5.7 show the predictions at the time they were made for each of the SSWs in each scenario. These are displayed as the thin lines, in red, blue and green for the three SSWs respectively. The observations made, if a human visited the SSW within the task time at the time of the prediction or not, is indicated by the thicker lines with the same color scheme. A “1” indicates a visit, and a “0” indicates no visit (the values are slightly shifted to avoid obscuration). An ideal prediction will be the same as the observation, it is desired to be as close to this as possible.

In the plot from scenario 1 we can see the apparent similarity between the predictions and observations made for SSW 2 and 3 (Figure 5.5). The most distinguishable difference is due to the shorter task time of SSW 2. The unevenness and peaks is most likely caused by the TBV, which is most apparent in SSW 2 and 3. The predictions are most of the time very close to one or zero.

In scenario 2 the observations are matching the peaks and valleys of the predictions quite well (Figure 5.6). The peaks and valleys are on the other hand not as high and deep in SSW 2. The short time between each visit makes the tails of the PDFs overlap between the visit. This turns out to be a challenge with this approach. Further, the effect of the TBV is only visible as small peaks in the curves. This both fits well with the impression of the well matching observations, but also might indicate that the TBV component of the prediction has little influence on the system, especially when the ToV component fits well.

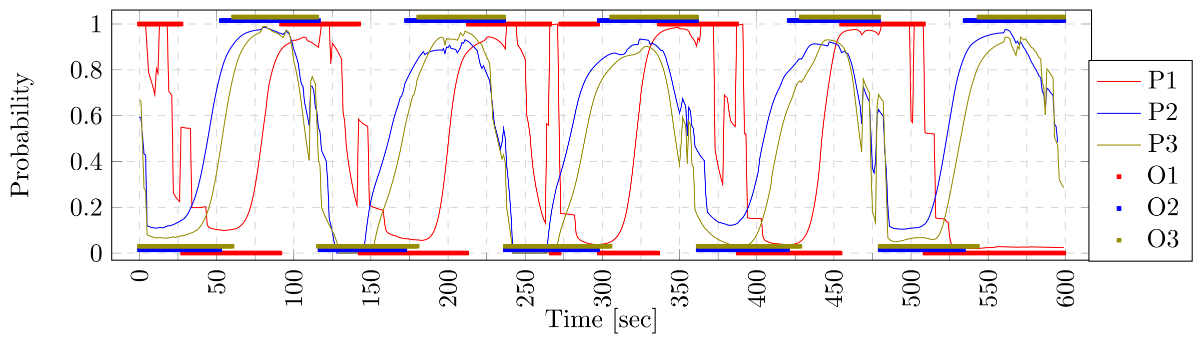


Figure 5.7: The predictions, P_i, through a day and the corresponding observations, O_i, for Scenario 3

The first difference seen in the plots from scenario 3 compared to the first two scenarios is that the system is now more affected by the TBV (Figure 5.7). The impact is especially dominant for $P1$ where sudden peaks can be seen. There is also much more overlap between the visitations, and times where no shared sub-workspace is visited. This corresponds well with the expectations from the scenario description. Also in this scenario, the plots indicate a good match between the predictions and observations.

5.3.1 Brier Score

The Brier Score (BS) is a proper score function that measures the accuracy of probabilistic predictions [116]. This is applicable to described scenarios as the outcomes are binary, a SSW is either visited by a human within the time frame or not. The score measures the mean square difference between the probability assigned to an outcome, and the actual outcome, $\mathbf{o}$. The score is thus always between zero and one, where a lower score indicates better predictions. A decomposed Brier Score provides insight on the deeper behavior of the binary classifier [115]. The three additive components; Reliability (REL), Resolution (RES) and Uncertainty (UNC) are used (5.1) and (5.2).

$$BS = REL - RES + UNC \tag{5.1}$$

$$BS = \frac{1}{N}\sum_{k=1}^{K} m_k(\mathbf{p_k} - \bar{\mathbf{o}}_\mathbf{k})^2 - \frac{1}{N}\sum_{k=1}^{K} m_k(\bar{\mathbf{o}}_\mathbf{k} - \bar{\mathbf{o}})^2 + \bar{\mathbf{o}}(1 - \bar{\mathbf{o}}) \tag{5.2}$$

The number of unique predictions are denoted K, this is the number of bins the predictions are organized in. In this paper the predictions are divided into 10% intervals, thus $K = 10$. With M being the number of predictions, $\bar{\mathbf{o}}$ is the base rate of events as shown in equation (5.3), m_k is the number of predictions in the same bin of K, and $\bar{\mathbf{o}}_\mathbf{k}$ is the observed frequency, given prediction $\mathbf{p_k}$.

$$\bar{\mathbf{o}} = \sum_{t=1}^{M} \frac{\mathbf{o}_t}{M} \tag{5.3}$$

The Reliability term measures how close the probabilities are to the true probabilities. This means that it is expected that the event will occur one out of ten times a 10% chance is predicted. The reliability term is the mean square of the predicted value and the expected value. A low reliability value contributes towards a low Brier Score. The Resolution term measures how much the observed frequencies differ from the base rate of events. A perfect resolution gives one, while the worst case gives zero. The uncertainty is the inherent uncertainty in the event. The most uncertain event is one that has a 50% occurrence rate. The uncertainty is zero if the event always or never occurs and 0.25 if the occurrence rate is 50%.

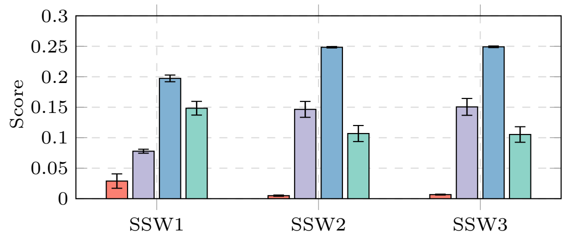
Further, the system is evaluated with the Brier Skill Score (BSS). The skill score is a measure of the difference between the score for the predictions and the score for an unskilled standard prediction. The range of the skill score is from $-\infty$ to 1, where all negative scores indicate a system that is less accurate than an unskilled standard prediction. A skill score of 1 indicates a prefect prediction model. The unskilled prediction used in this paper is the base rate referred to earlier, which gives us the BSS as shown in (5.4)

$$BSS = 1 - \frac{BS}{BS_{ref}} = \frac{RES - REL}{UNC} \tag{5.4}$$

In Appendix B the development of the different scoring components is shown. The effect of the relearning is not too apparent, even over ten days. Most of the values are in the same order of magnitude as that of the first day. This, however, might also indicate good performance after the first day. The uncertainty values (UNC) for SSW2 and SSW3 are 0.2484 and 0.2492 respectively and indicate that each of the zones are occupied close to 50% of the time. This level of uncertainty describes a challenging event to predict. The uncertainty in SSW1 is still high, however significantly lower than that of SSW2 and SSW3. The reliability is on the other hand very good for all three tasks. The DBS for the three cases are not as low as one might wish for, however, the values of the different components can be compared further. A high uncertainty is mainly compensated for with a high resolution. The resolution is, however, not high enough to deal with the very high uncertainty. The best resolutions can be seen with SSW2 and SSW3 with 0.1463 and 0.1506 respectively, on the other hand the uncertainty is equally higher in these SSWs. The Brier Score is the best for SSW2 and SSW3 with 0.1069 and 0.1053 respectively, although not by much.

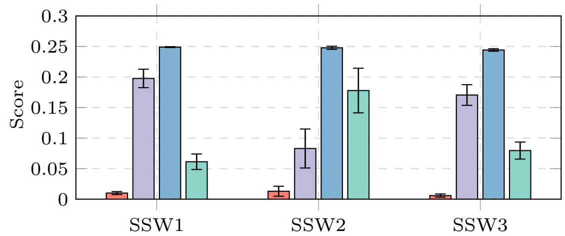
Studying the skill scores reveals a greater difference in the predictions between SSW1 and SSW2 and SSW3 (Figure 5.8). The skill score the system achieved for SSW2 and SSW3 are more than double that for SSW1, almost triple. The scores of 0.57 and 0.58 describe a predictor that is closer to perfect than to a base rate predictor.

Similarly to scenario 1, scenario 2 is dealing with events with a very high uncertainty, more than 0.24 for all SSWs (Figure 5.9). Again, this indicates the challenge in predicting the event. SSW1 and SSW3 have very high skill scores, these were the assembly areas and were visited three times each during the work day. This makes them easier to predict than the pick up areas that were visited for a shorter time, six times a day. The skill score for SSW2 is not even half of that of SSW1 and SSW3. The resolution is similarly to scenario 1 not high enough to deal with the very high uncertainty.



REL		0.0289	0.0049	0.0067
RES		0.0778	0.1465	0.1506
UNC		0.1973	0.2484	0.2492
DBS		0.1484	0.1069	0.1053

Figure 5.8: Average Decomposed Brier Scores for Scenario 1



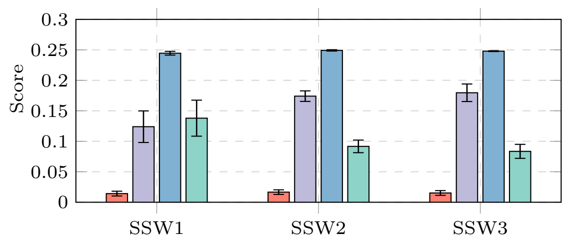
REL		0.0100	0.0130	0.0059
RES		0.1977	0.0830	0.1705
UNC		0.2490	0.2478	0.2442
DBS		0.0613	0.1779	0.0796

Figure 5.9: Average Decomposed Brier Scores for Scenario 2

The scores from scenario 3 show many of the same traits as that of scenario 1 and 2 (Figure 5.10). There is a very high uncertainty, a good reliability and a resolution that is not high enough to give a very good Brier Score or Skill score. The Brier Score and Skill score are on the other hand rather good, especially for SSW2 and SSW3, where the operator spent the most time.

5.3.2 Total Scores

The combined DBS and BSS for all experiments can now be calculated to investigate whether or not hypothesis $\mathbf{H_{1.1}}$ holds. Over the course of 3 scenarios, all with 3 SSWs and 600 predictions



REL		0.0174	0.0166	0.0152
RES		0.1240	0.1741	0.1796
UNC		0.2445	0.2491	0.2479
DBS		0.1379	0.0917	0.0835

Figure 5.10: Average Decomposed Brier Scores for Scenario 3

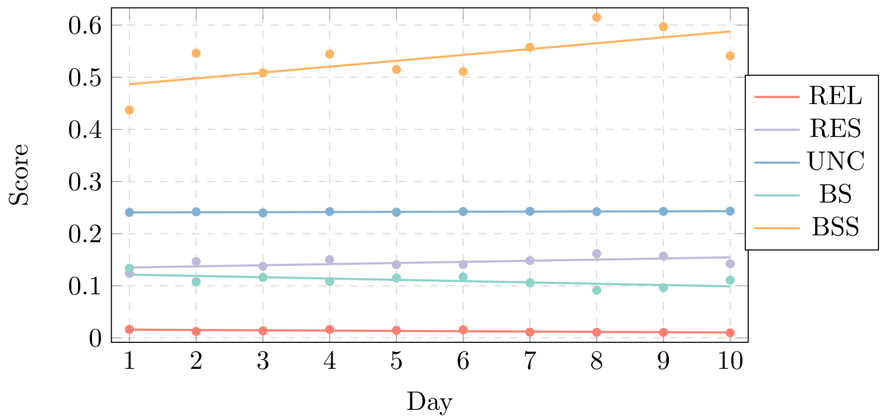


Figure 5.11: Scores for all predictions by day.

Table 5.8: The average REL, RES, UNC, DBS and BSS for all predictions.

	REL	RES	UNC	DBS	BSS
Avg.	0.0132	0.1449	0.2419	0.1103	0.5373
Std.	0.0085	0.1840	0.2488	0.0732	0.7056
Slope	-0.0006	0.0022	0.0003	-0.0025	0.0113
MSE	0.0013	0.0068	0.0006	0.0071	0.0291
Improvement	34.1%	14.4%	0.99%	18.5%	20.8%

per SSW, each ran for 10 workdays, the total number of predictions can be calculated. The total number of predictions throughout the experiment is then 3 scenarios $\times 3$ SSWs $\times 600$ iterations $\times 10$ days = 54000. The average REL, RES, UNC, DBS and BSS for these 54000 predictions are shown in Table 5.8. The system has a DBS of 0.1103 and a BSS at 0.5373. The BSS indicates that the system is slightly closer to a perfect predictor than a base rate predictor. The low DBS also indicates a well performing predictor and for the purpose of further development it is concluded that $\mathbf{H_{1.1}}$ holds.

The effect of the relearning is investigated to confirm hypothesis $\mathbf{H_{1.2}}$ by combining all predictions made for each of the days. The resulting plot can be seen in Figure 5.11. A positive development would mean an increasing RES and BSS, and a decreasing REL, UNC and DBS. Using linear regression the slope of the development of the scores can be found. The slope and the MSE from the regression can be seen in Table 5.8. Although every score has the desired sign, the slope for both UNC and REL is too small to indicate any real change. Although the REL improves with 34.1% over 10 days, it is already very good from day 1, so the effect of the improvement is minimal. The improvement of UNC is less than 1%, however, it was not expected to improve, as it is based on the observations alone, and not the predictions. The slope of RES and BS are also small but indicate a positive development as more data is gathered. Their improvements are 14.4% and 18.5% respectively. The most significant development can be seen in the skill score, BSS. Being a result of the slight improvement of all other scores, it is as expected. An improvement of 20.8% over 10 days clearly indicates the effect of the continuously relearning model in the system, $\mathbf{H_{1.2}}$ thus holds.

5.4 Summary

The purpose of these experiments was to investigate how the system corresponds to **PS1**. The proactive abilities of the system is directly related to the performance of the likelihood analysis.

PS1: *The developed system should act proactive against dangers.* Today's safety systems moves the robot away if it is in a conflict with the human to avoid a collision. The developed system should avoid human-robot conflicts, thus acting as a new layer of safety.

Two hypotheses were presented to verify **PS1**, firstly that the system is able to predict whether the human operator will occupy a given SSW within the robot's task execution time. Secondly, that the system's predictive capabilities would improve over time, as more observations were made, and the Computation/Learning component was given more time to calculate appropriate parameters. This was verified using evaluations of the results of DBS and BSS.

The total Brier Score over 54 000 predictions at 0.1103 and Brier Skill Score at 0.5373 clearly indicates the system's ability to predict whether a SSW will be occupied by a human within the robot task execution time, or not. It is therefore safe to conclude that hypothesis $\mathbf{H_{1.1}}$ holds. Over the course of 10 "workdays" the systems skill improved by 20.8% and the Brier score for the system improved by 18.5%. This demonstrates the system's ability to adjust, relearn, and improve as more data is gathered, and hypothesis $\mathbf{H_{1.2}}$ thus holds. The performance of the likelihood analysis is therefore satisfactory for a proactive system, and thus far can it be said that **PS1** holds. The next chapter will experimentally test the proactive decisions of the system further.

Chapter 6

Performance of the Responsible Robot Based HRC

6.1 Introduction

This chapter will describe the experiments conducted to verify the system's compliance with **PS1** and **PS2**. These involve the proactiveness of the system and its ability to maintain productivity. An experiment was set up to compare the performance of the proposed system to that of a preprogrammed robot assistant. The experiments were conducted with the decisions in the system based solely on Level 3 SA. This might not eliminate all human-robot conflicts, however, the goal of this new safety layer is to reduce the number of human-robot conflicts as much as possible. A safety layer with a reactive response to any conflict would be necessary to realize a proper safety system. However, the purpose of these experiments is to test the effects of Level 3 SA, and the Level 2 SA will not influence the decisions not to confuse which level of SA is causing the decision.

6.2 Experiments

An assembly task was devised where the human and the robot each had a set of tasks. The participants then completed one "workday" as a human-human collaboration (HHC) as a reference to the later experiments. The participants then completed two more "workdays", one as a preprogrammed human-robot collaboration (PP-HRC) and one as a Responsible Robots based human-robot collaboration.

6.2.1 Experimental Setup

The experimental setup was built around the same setup as in Chapter 5, a workspace was built around a desk. A Kinect V2 sensor was used to track the human operator. The code was developed in LabVIEW, with the HARO3D™[1] VI library for integration with the Kinect.

[1] HaroTek LCC. www.harotek.com (Accessed 11/10/2015)

Two off-the-shelf desktop PCs were used, one of which was dedicated to handle the Computation/Learning component. The desk was located between the human and the robot, allowing the human and the robot to work on the desk from each side of the table. The desk used measured 180 cm by 40 cm. The workspace was separated into human sub-workspaces (HSWs) and robot sub-workspaces (RSWs) based on their tasks. The areas where a HSW and a RSW are overlapping are the shared sub-workspaces (SSWs) as shown in Figure 6.1. Note that HSWs, RSWs and thus also the SSWs might also overlap themselves. The HSWs were fixed for all scenarios and defined as three equal parts of the desk, HSW1 being the leftmost part of the desk, HSW2 the middle and HSW3 the rightmost for the human. The RSWs were defined from the robot's tasks in the different scenarios. A second table was placed on the robot's side serving as a storage place for full and empty brick containers.

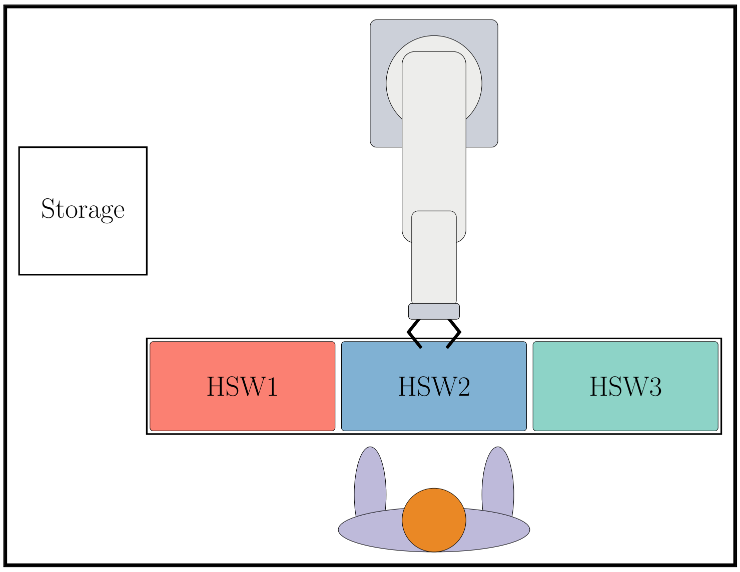


Figure 6.1: Layout of the workcell used in the experiments.

6.2.2 Procedure

When the participant first arrived at the work space, they received training from an instructor. They then got a few minutes to familiarize themselves with the tasks, until they were comfortable enough to proceed. They then went on with completing a work day, while being observed by the system. The participants had no assistant through this day and they had to bring the boxes with bricks back themselves. This workday was used as day zero, and the observations were stored in the system's Memory component while the Computation/Learning component processed them. The participants then worked one workday while collaborating with the instructor as a human-human collaboration (HHC). The instructors task were the same as the robot would later have. The participants then worked one workday as a preprogrammed human-robot collaboration (PP-HRC), or a Responsible Robots based HRC. The PP-HRC included a robot that had been preprogrammed to perform its tasks at given points in time based on a predefined work pattern. A workday with the final approach, Responsible robot based HRC or PP-HRC was then performed. A summary of the steps is shown in Table 6.1. All the trials were recorded, and the video material was reviewed to gather the relevant data.

Table 6.1: Summary of the procedure the participants went through during an experiment.

Step	Procedure
1	Receive training
2	Work one day on their own
3	Work one day as HHC
4	Work one day as PP-HRC/Responsible Robot based-HRC
5	Work one day as Responsible Robot based-HRC/PP-HRC

6.2.3 Scenario

The system was tested in a scenario where the human was building two Lego figures at a time, then taking them a part (Table 6.2). The instructions had a low level of detail which allowed for a variety of different ways to complete the tasks in the instructions. The human operator could choose how to complete the tasks, and also vary how the tasks were completed throughout operation. The robot's tasks involved replenishing bricks and containers, and picking up full containers (Table 6.3 and resulted in three SSWs, two of which were overlapping (Figure 6.2)). The robot's task execution times were set as the time it would need from it decides to start the task, until it is clear of the SSW. Because of this, the robot uses more time than the task execution time before it is ready to start a new task. In this scenario the human would be building the figures in a sub-workspace not shared with the robot, thus leaving all the SSWs unvisited at the same time. The subjects were able to finish anything between three and five cycles during one workday.

Table 6.2: Human work instructions

Step	Action
1	Pick 12 large and 24 small bricks of any color in HSW1
2	Pack bricks in an empty box
3	Bring the box with bricks and assemble two figures in HSW2
4	Bring the figures to HSW3 and take them apart
5	Pack used Lego bricks in an empty box
6	Repeat step 1-5 until end of day

6.2.4 Model Components

The design of some of the components in the proposed model for Responsible Robots based-HRC depend on the implementation. These include Hazard Identification, Goals and Objectives, and the Decision components, these will be described in this section.

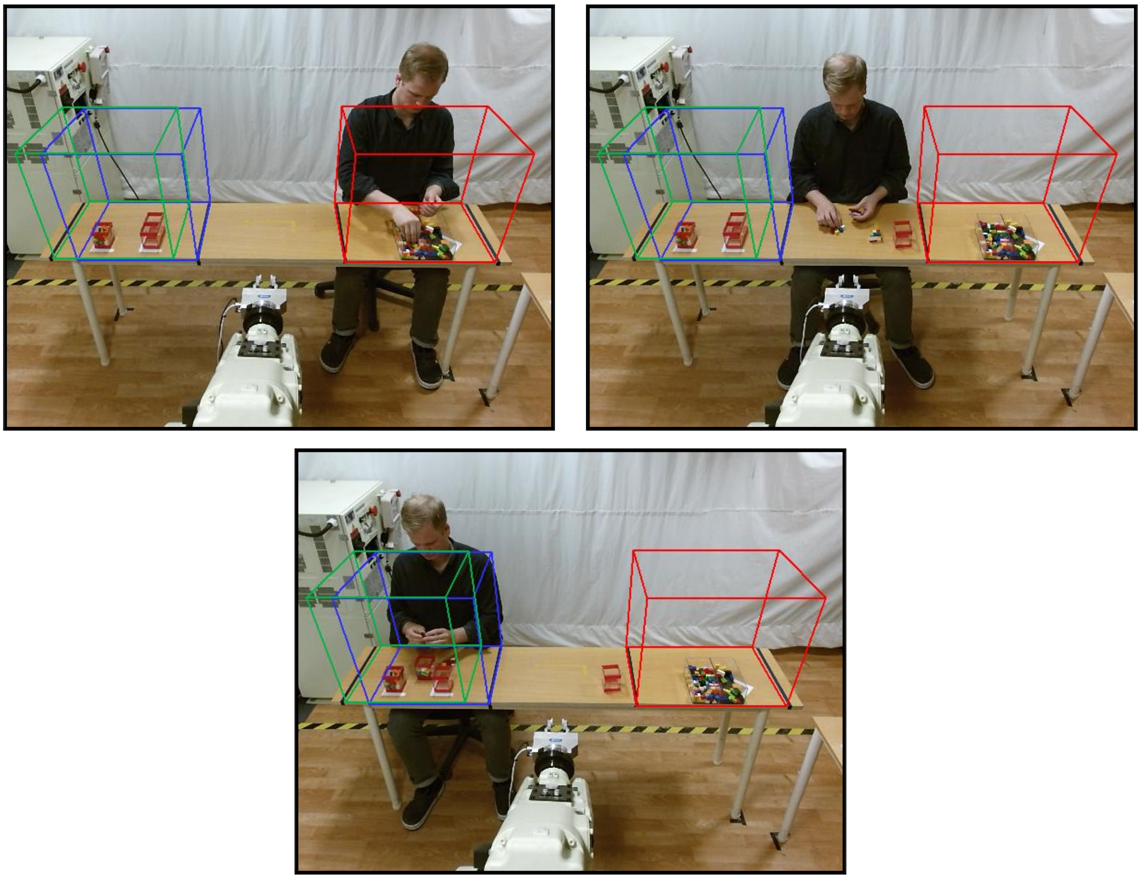


Figure 6.2: The shared workspace for the scenario, with the three shared sub-workspaces indicated by red, blue and green respectively.

Table 6.3: Robot Tasks

Task	Action	Time
1	Replenish bricks in brick container in HSW1	25
2	Replenish empty containers to HSW3	24
3	Collect container with used bricks from HSW3	18

Hazard Identification

As discussed in Section 4.2 the most essential hazards in a HRC are any form of unwanted contact between the human and the robot. This is therefore what will be regarded as the main hazard in these experiments and the only hazard included, that the human and the robot is occupying the same space at the same time.

Goals and Objectives

The primary objective was to keep the risk below a level of 0.5. Further, the system had a goal to complete as many tasks as possible, with an objective to not exceed an imbalance of one task being executed more than two times more than the others. The task progression was monitored and tasks with fewer executions were thus prioritized.

Decision

To investigate the contribution of the proactive safety layer, there is no other safety layers included in the system. Basic necessary safety measures such as emergency stop is included externally. The decision making mechanism in the system is thus only affected by the risk related objective and the productivity objective. The algorithm governing the decision component is depicted in Figure 2. Essentially, the system has four available options: execute Robot Task 1, Robot Task 2 or Robot Task 3, or it can wait. First, the tasks with a risk lower than the set threshold is identified. If no task have a low enough risk associated with it, the system waits. The number of executions of each task is then compared, and any task that has been executed two times or more than the least executed task is excluded from selection. The task with the combination of the lowest risk, and the lowest number of executions is then selected for execution. The system continuously runs this decision loop as long as the robot is in stand by. A system with all three safety layers would continuously run the decision loop, and decisions based on low level SA would overrule and interrupt actions started on the basis of higher level SA.

Algorithm 2 Decision Component

procedure SELECT ACTION
    if Robot busy = FALSE **then**
        if Risk(Task$_i$) $<$Risk$_{max}$ **then**
            if Task Penalty(Task$_i$) $<$Task Penalty$_{max}$ **then**
                Start Task$_i$
                Task penalty(Task$_i$) $\leftarrow$ Task Penalty(Task$_i$)+increment
            else Return
        else Return
    else Return

6.2.5 Parameter Settings

The workday was set to 10 minutes, and the maximum TBV set to 1 minute. Both the ToV and the TBV dataset were distributed over 200 bins, giving a time interval of 3 seconds and 0.3 seconds in (4.10) and (4.11) respectively. The maximum accepted MSE for the fitness of the curve fitting was set to $\varepsilon_{max} = 5 \times 10^{-4}$. The maximum number of PDFs was set to 5 and the relearn rate was set to 25%.

The design parameters used in the human motion predictions, k_1, k_2 and γ, were set to $k_1 = 1$, $k_2 = 0.25$ and $\gamma = 1$. The length of the human velocity prediction history, L, was set to $L = 1 second$. The maximum distance of the proximity field ρ was set to $\rho = 20cm$. The likelihood analysis iteration time was set to 1 second.

6.2.6 Participants

The experiments were conducted with a total of 9 participants. The participants were students at Chuo University and were all male in their 20's. The participants were recruited by email and word-of-mouth.

6.2.7 Hypothesis

A set of hypotheses is formulated to help determine whether or not the system is in compliance with **PS1** and **PS2**. The forecaster used to make proactive decisions was thoroughly tested and verified in Chapter 5, in this chapter the system's ability to make safe proactive decisions in a real human-robot collaboration will be tested. The three indicators were used for verification of the two **PS**s.

First of these indicators is the precision, which is the rate at which true positives (TP) occur, in a set of positives. A positive is given if the system estimates the risk to be low enough, and it is safe to start a task. A set of positives is thus all the positives throughout a workday. A false positive (FP) is when it turns out that it was not safe, and the human and the robot enters the same SSW at the same time. The precision is the probability that it is safe to start a task, given the system decides to start a task.

Secondly, it is important that the robot manages to keep up with the operator's work pace. To keep up the robot was in this scenario required to completed 3 tasks per cycle completed by the operator. If the robot cannot keep up with the human's pace, it might result in frustration and loss of concentration for the human. No one likes to wait and be delayed by coworkers that cannot keep up with one's pace.

Lastly, the number of human-robot conflicts was counted. A conflict is when the human and the robot is present in the same SSW at the same time. These conflicts might not pose a danger for the human, however it is frustrating and stressful to work with anyone who continuously interrupts one's work. Avoiding these conflicts might lead to a much more comfortable and relaxed work situation for the human. This is important to allow the human to focus more on his/her task, rather than where the robot is and what it will do next. The safety in a conflict would be resolved by a reactive Level 2 SA system. However, in these experiments there is no reactive layer to get a clear picture of the proactive layer's performance.

The goal of the system is therefore to reduce the number of human-robot conflicts, have a high precision and being able to produce at the same, or higher, rate as the human.

$\mathbf{H_{2.1}}$ The system has a high precision (rate of true positives).

$\mathbf{H_{2.2}}$ The system reduces the number of human robot conflicts.

$\mathbf{H_{2.3}}$ The system is able to keep up with the human operators productivity.

It can safely be concluded that the system is in compliance with the two statements, **PS1** and **PS2**, if these three hypothesis holds in combination with the results from Chapter 5.

6.3 Results

The experiments were conducted as previously described with a total of 9 participants. The HHC case was always conducted first, to give the system time to learn, then either the PP-HRC or Responsible Robots based-HRC were second. A total of 5 participants followed the HHC→PP-HRC→Responsible Robots based-HRC pattern, and 4 followed the HHC→ Responsible Robots based-HRC→PP-HRC pattern. Due to practical reasons the learning time allowed to the system was limited to a minimum, and the Responsible Robots based-HRC case was conducted after the first or second iteration, typically.

The human task was as expected solved in a variety of ways. Step 1 in Table 5.6 was by some solved by counting a few bricks at the time into one hand, then dropping them in the box. Some counted the bricks one by one using one hand, others with both hands. When building the figures some built the two figures step by step in parallel and some built one at the time. Some of the participants built the figure on the table, while some held the figure in one hand while building. Similarly when taking the bricks apart, some took them both apart on the table, then scooped the bricks into the box, some picked the bricks directly of the figure and into the box and some broke the figures into smaller pieces on the table, then took those apart directly into the box. These variations, and many other minor variations were observed. The different approaches to solve the task also led to a variation in the relative time spent on each sub task. Some spent the most time on building while the others spent more time on picking apart than building.

During the experiments the necessary data to evaluate the system based on the performance indicators was gathered (see Table 6.4). The data was found by carefully going through video recordings of the experiments. Note that the required number of Robot Tasks are not necessarily an integer. The required number is calculated based on the progression of the human operator which may have been interrupted in the middle of a task cycle at the end of the work day.

Table 6.4: Data gathered during the experiments regarding system performance. (Responsible Robots based-HRC is denoted RR-HRC)

	Positives		Human-Robot conflicts		Robot Tasks	
#	TP	FP	PP-HRC	RR-HRC	Required	Completed
1	12	1	5	1	10.5	13
2	13	1	3	1	9.9	14
3	11	1	3	1	6.6	12
4	13	0	0	0	12	13
5	15	0	6	0	16.2	15
6	10	2	1	2	12.9	12
7	13	0	5	0	13.5	13
8	14	0	0	0	12.6	14
9	10	0	3	0	9.6	10

Table 6.4 shows that the system considered the situation safe and decided to start a robot task a total of 116 times throughout all 9 experiments, as shown in Table 6.5. Amongst these 111

were completed successfully without causing a human-robot conflict, while 5 caused a conflict. This gives an average precision of $\frac{TP}{TP+FP} = \frac{111}{111+5} = 0.96$.

Table 6.5: Number of true and false positive predictions.

Positives	TP	FP	Precision
116	111	5	96%

Further, another important performance indicator is the reduction of human-robot conflicts. In the PP-HRC case the human alone was responsible for avoiding conflicts, while it was a shared responsibility with the Responsible Robots based-HRC approach. In Table 6.4 can it be seen that there in most cases is a great reduction of conflicts with the Responsible Robots based-HRC approach. In only one case (participant 6) was there one more conflict with the proposed approach. Both participants 4 and 8 avoided conflicts totally with both methods. The mean reduction of conflicts is then 2.3 fewer conflicts through the workday in this experiment (Table 6.6). The best improvement was experienced by participant 5 who went from 6 conflicts to none.

Table 6.6: Difference in number of human robot conflicts between methods

	RR-HRC - PP-HRC
Mean	-2.3
SEM	0.8
SD	2.4
Minimum	-6
Maximum	1
Count	9

The productivity rate was investigated as the last performance indicator. The robot was required to complete three tasks for each human work cycle to be able to keep up. The required number of robot tasks as shown in Table 6.4 was therefore found by multiplying the number of completed work cycles by the human operator with three. Compared to the number of completed tasks, it becomes apparent that the robot is ahead in 6 of the 9 experiments. For participants 6 and 7 the robot was less than one task behind. The robot was more than one task behind for participant 5 only, whose pace was by far the fastest in the group. This pace was borderline on the capabilities of the system with the selected speed of the robot. The mean difference between required and completed robot tasks is shown in Table 6.7.

6.4 Evaluation of Results

A statistical analysis of the data was carried out by comparing the means with a Student's t-test. Independent samples or paired samples were used depending on the comparison. The alpha

Table 6.7: Number of robot tasks completed compared to the pace of the human

	#Completed Tasks - #Required Tasks
Mean	1.4
SEM	0.8
SD	2.3
Minimum	-1.2
Maximum	5.4
Count	9

level was set arbitrarily to 0.1 and the null hypothesis was that the means are equal for the paired or equal to 0 for the individual samples.

As described in Section 6.2.7 there were some goals with implementing the Responsible Robots based-HRC approach and some hypotheses were presented. The first hypothesis, $\mathbf{H_{2.1}}$, states that the system has a low rate of false positives. Therefore, the precision was, as previously calculated, 0.96 which is equivalent to a False Positive Rate (FPR) of 0.04. In other words, the probability that there will be no conflict if the system starts a task is 0.96. The conflicts that occurred in the conducted experiments yielded in general low risk and would be absorbed by a reactive layer. The speeds were low at all times, and there was no use of hazardous tools. A different scenario with potentially greater consequences might result in a different precision. The data is evaluated with a one tailed Poisson confidence interval at 95% (Table 6.8). This yields a Precision greater than 91% with a 95%CI and it is therefore concluded that $\mathbf{H_{2.1}}$ holds.

Table 6.8: Number of true and false positive predictions.

Positives	TP	FP	Precision 95%CI
116	111	5	> 91%

Further, the reduction of human-robot conflicts was tested in accordance with $\mathbf{H_{2.2}}$. The mean of the reductions was 2.3, for the difference to be significant the alternative hypothesis that the mean is less than 0 must be true. An individual sampled one-tailed t-test gives

$$\begin{aligned} t(9) &= 2.92, \\ p &= 0.0097, \\ p < \alpha &\rightarrow \text{null hypothesis rejected}, \\ H_a : \Delta_{mean} < 0 &\rightarrow \text{is true}. \end{aligned}$$

Resulting in the conclusion that the reduction of human-robot conflicts is significant and that $\mathbf{H_{2.2}}$ holds.

The data regarding the pace of the robot was analyzed as a last indicator to verify the performance of the Responsible Robots based-HRC approach. In this case the number of tasks completed by the robot should exceed the required number based on the human's pace. The alternative hypothesis is therefore that the mean of the difference is greater than 0. An individual sampled one-tailed t-test gives

$$
\begin{aligned}
t(9) &= 1.79, \\
p &= 0.056, \\
p < \alpha &\rightarrow \text{null hypothesis rejected,} \\
H_a : \Delta_{mean} > 0 &\rightarrow \text{is true,}
\end{aligned}
$$

which also in this case concludes with a significant difference. The Responsible Robots based-HRC approach is therefore capable of keeping up with the human operator and $\mathbf{H_{2.3}}$ thus holds.

All three indicators used to evaluate the system performance resulted in favor of the proposed system and all three hypotheses were kept. It is therefore concluded that the performance of the Responsible Robots based-HRC system is as expected.

6.5 Summary

The purpose of these experiments was to verify if the proposed system is in compliance with **PS1** and **PS2**.

PS1: *The developed system should act proactive against dangers.* Today's safety systems moves the robot away if it is in a conflict with the human to avoid a collision. The developed system should avoid human-robot conflicts, thus acting as a new layer of safety.

PS2: *The developed system should be able to solve the necessary tasks to maintain its productivity.* The system should be designed to be independent of task and robotic hardware. Further, the developed system should have an awareness of what the human operator expects of it.

The performance of the proposed Responsible Robots based HRC was experimentally tested with several test subjects. The participants worked alongside a responsible robot on an assembly task, and the robotic system decided autonomously when to safely execute its own tasks and when to wait. Three hypotheses were formulated to investigate this performance. Firstly, that the system had a low rate of false positives. Secondly that the system reduced the number of human robot conflicts. And lastly that the system is able to keep up with the human operator's productivity. An alpha level of 0.1 was arbitrarily chosen for evaluation purposes.

Firstly, the rate of false positives was found to be 0.04, which is well below the set alpha value. It was therefore concluded that the system has a low rate of false positive and hypothesis $\mathbf{H_{2.1}}$ holds. Further, the very important hypothesis on the reduction of human robot conflicts was investigated. The average reduction during the 10 minute workday was 2.3 and the difference was confirmed to be significant by using a student-t test. It is therefore safe to say that the proposed method reduces the number of human robot conflicts and $\mathbf{H_{2.2}}$ holds. Lastly, the last hypothesis about the productivity of the system was investigated. It was found that the system on average was 1.4 tasks ahead of the human, and the difference was confirmed to be significant in a student-t test. Therefore, it is concluded that $\mathbf{H_{2.3}}$ also holds.

The experiments verified the performance and demonstrated a reduced number of human robot conflicts, and an ability to keep up with the human's work pace, and its precision in decision making. It is therefore safe to conclude that the system's performance corresponds to the statements **PS1** and **PS2**.

Chapter 7

Effects of working with a Responsible Robot in a HRC

7.1 Introduction

In this chapter, the experiments related to **PS3** will be presented. It is stated that the proposed system should be designed to improve the effect the collaboration has on the human operator. It is vital that the human operator is able to maintain focus on his/her task to have a fruitful collaboration. If the human operator is continuously interrupted by the robot's unpredictable movements, it might cause stress and unnecessary time pressure and frustration. These experiments will investigate how the proposed system affects this.

7.2 Experiments

A series of experiments with several human participants was conducted. The purpose was to explore the differences in stress and perceived workload between collaborating with a human, a pre-programmed robot and a RR. An assembly task was devised where the human and the robot each had a set of tasks. The human operator's instructions had a low level of detail which allowed for a variety of different ways to complete the tasks in the instructions. The human operator could choose how to complete the tasks, and also vary how the tasks were completed throughout operation. This is important as it provides a more natural behavior by the human operator. The participants then completed one "workday" as a HHC, one in collaboration with a preprogrammed robot and one working with a Responsible Robot.

7.2.1 Experimental Setup

The same work cell and system that was as used in Chapter 6 was also used in this experiment. A Kinect V2 sensor was used to track the human operator. The code was developed in LabVIEW, with the HARO3D™[1] VI library for integration with the Kinect. Two off-the-shelf desktop PCs

[1] HaroTek LCC. www.harotek.com (Accessed 11/10/2015)

were used, one of which was dedicated to do the curve fittings. The desk was located between the human and the robot, allowing the human and the robot to work on the desk from each side of the table. The desk used measured 180 cm by 40 cm. The workspace was separated into human sub-workspaces (HSWs) and robot sub-workspaces (RSWs) based on their tasks. The areas where a HSW and a RSW overlap are the shared sub-workspaces (SSWs) as depicted in Figure 7.1. Note that HSWs, RSWs and thus also the SSWs might also overlap themselves. The HSWs were defined as three equal parts of the desk, HSW1 being the leftmost part of the desk, HSW2 the middle and HSW3 the rightmost for the human. A second table was placed on the robots side serving as a storage place for full and empty brick containers.

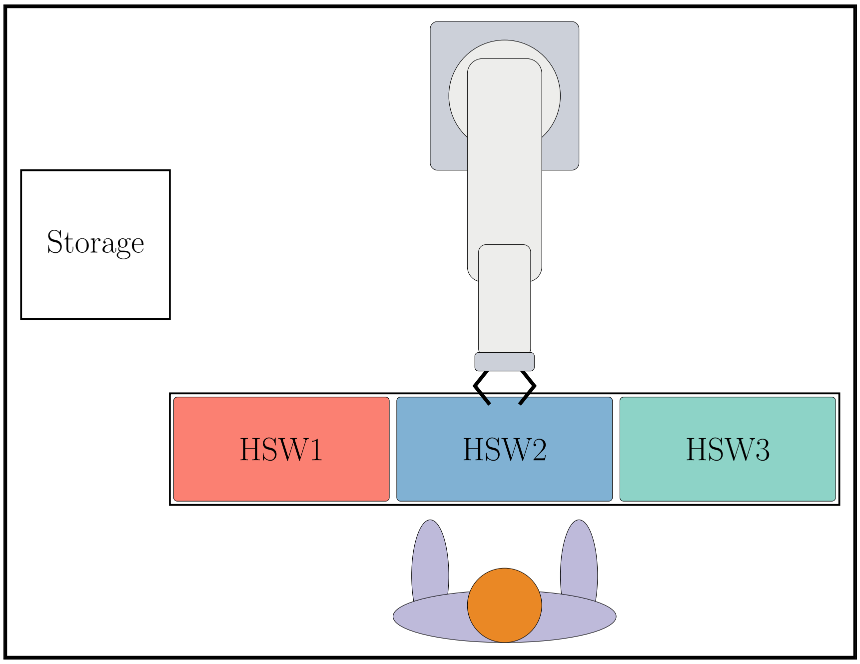


Figure 7.1: The layout of the collaborative work cell used in the experiments.

7.2.2 Procedure

When the participant first arrived at the work space, they received training from an instructor (Table 7.1). They then got a few minutes to familiarize themselves with the tasks, until they were comfortable enough to proceed. They then went on with completing a work day, while being observed by the system, recording the observations to the memory component. The system Computations/Learning component immediately starts to calculate proper parameters for the likelihood analysis. The participants had no assistant through this day and they had to bring the boxes with bricks back themselves. This workday was used as day zero, and the learning algorithms were applied to the data. The participants were then instructed to read about the different scales in the NASA-Task Load Index test (See Section 7.2.4), and arrange them by importance. The scale that was most important to them was to be assigned a weight of 6, the least important a 1. The participants then worked one workday while collaborating with the instructor. The instructor's tasks were the same as the robot would later have. A NASA-TLX form was then filled out for the HHC case. The participants then went on with the PP-HRC case, or the Responsible Robots based-HRC case. Another NASA-TLX form was filled out for the performed case. A workday with the final approach, Responsible Robots based-HRC or PP-HRC was then performed before the final NASA-TLX form was filled out. A summary of the steps is shown in Table 7.1.

Table 7.1: Summary of the procedure the participants went through during an experiment.

Step	Procedure
1	Receive training
2	Work one day on their own
3	Assign weights to NASA-TLX scales
4	Work one day as HHC
5	Fill out a NASA-TLX form
6	Work one day as PP-HRC/Responsible Robots based-HRC
7	Fill out a NASA-TLX form
8	Work one day as Responsible Robots based-HRC/PP-HRC
9	Fill out a NASA-TLX form

7.2.3 Scenario

The same scenario that was used in the performance experiments (Chapter 6) was also used in these experiments The system was tested in a scenario where the human was building two Lego figures at a time, then taking them apart (Table 7.2). The robot's tasks involved replenishing bricks and containers, and picking up full containers (Table 7.3) and resulted in three SSWs, two of which were overlapping (Figure 7.2). The robot's task execution times were set as the time it would need from it decides to start the task, until it is clear of the SSW. Because of this, the robot uses more time than the task execution time before it is ready to start a new task. In this scenario the human would be building the figures in a sub-workspace not shared with the robot, thus leaving all the SSWs unvisited at the same time. The subjects were able to finish anything between three and five cycles during one workday.

Table 7.2: Human work instructions

Step	Action
1	Pick 12 large and 24 small bricks of any color in HSW1
2	Pack bricks in an empty box
3	Bring the box with bricks and assemble two figures in HSW2
4	Bring the figures to HSW3 and take them apart
5	Pack used Lego bricks in an empty box
6	Repeat step 1-5 until end of day

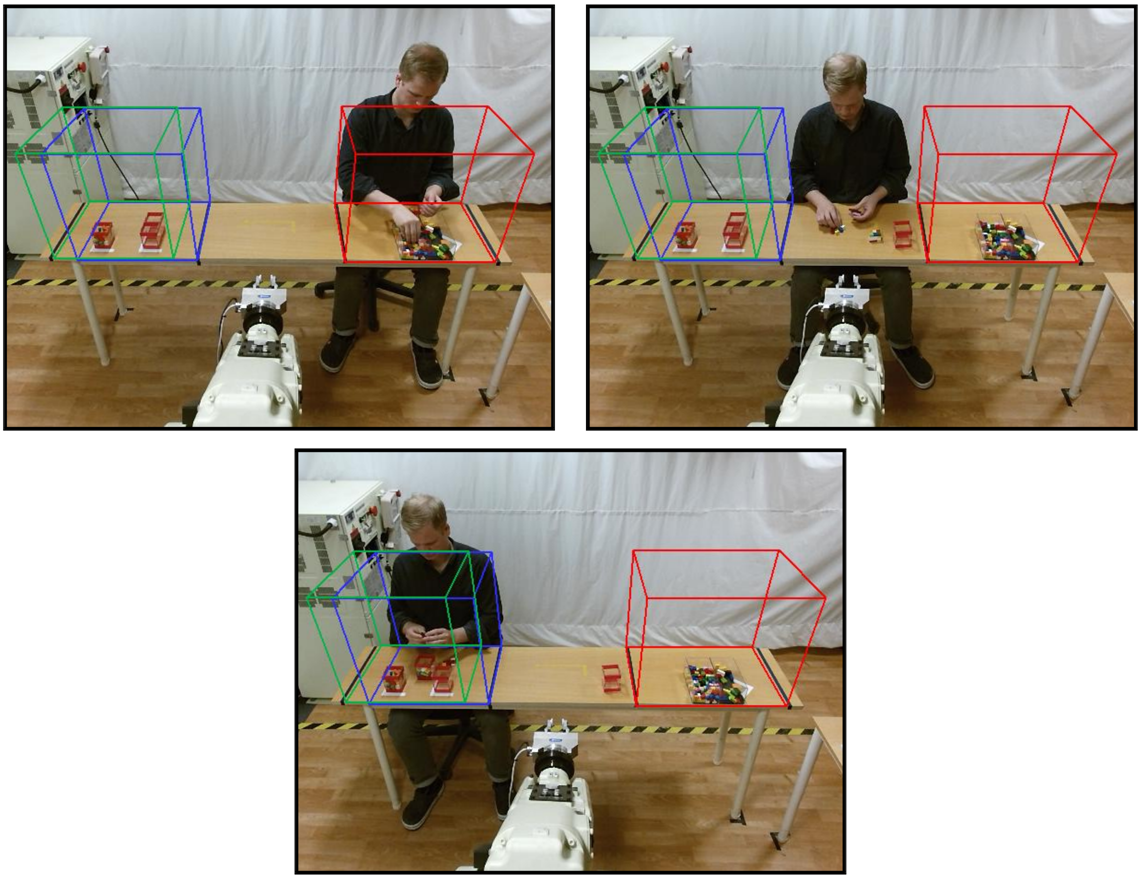


Figure 7.2: The shared workspace for the scenario, with the three shared sub-workspaces indicated by red, blue and green respectively.

Table 7.3: Robot Tasks

Task	Action	Time
1	Replenish bricks in brick container in HSW1	25
2	Replenish empty containers to HSW3	24
3	Collect container with used bricks from HSW3	18

7.2.4 NASA Task Load Index

In these experiments the participants were asked to complete a NASA Task Load Index form (NASA-TLX) after each test to test the perceived workload of the task [117]. The scales Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Performance (PE), Effort (EF) and Frustration (FR) were used (Table 7.4). Each of the scales were given a value between 0 and 100, where 0 denotes no load at all while 100 is perceived as a maximal load.

The participants were asked to weigh the different scales before the experiments started which were used to calculate a user weighted workload (UWWL). While the ambition of the system was not necessarily to reduce all the scales, a set of weights was also chosen by the research team to calculate a research weighted workload (RWWL). As shown in table 7.5, the

Table 7.4: The scales used in the NASA-TLX

Scale	Description
Mental Demand	The amount of mental activity that was required. How much thinking, deciding, remembering, searching etc
Physical Demand	How much physical activity was required
Temporal Demand	The amount of time pressure felt due to the rate at which the task elements occurred
Performance	How successfully was the task completed. How satisfied were you with your performance
Effort	How hard did you have to work to accomplish your level of performance
Frustration	How irritated, stressed, and annoyed were you, compared to calm and relaxed

scores with the highest priority to reduce are Temporal Demand and Frustration, and the lowest weights are given to Performance and Physical Demand.

Table 7.5: Selected weight for the TLX scales by the research team.

Weight	Scale
6	Temporal Demand
5	Frustration
4	Effort
3	Mental Demand
2	Performance
1	Physical Demand

The UWWL and RWWL are calculated as the sum of the multiples of the weights and scores given to each sub scale. The two scores UWWL and RWWL are calculated with two sets of weights set by the participant and the researcher respectively. The scores are then calculated as shown in (7.1) and (7.2) respectively, where s_i is the score denoted by the participant to sub scale i. The weights set by the user and researcher are denoted as w_{ui} and w_{ri} respectively. A lower UWWL or RWWL indicates a lower workload, and is thus the goal of the system.

$$UWWL = \frac{1}{\sum_i w_{ui}} \sum_i w_{ui} s_i \tag{7.1}$$

$$RWWL = \frac{1}{\sum_i w_{ri}} \sum_i w_{ri} s_i \tag{7.2}$$

7.2.5 Model Components

The design of some of the components in the proposed model for Responsible Robots based-HRC depend on the implementation. These include Hazard Identification, Goals and Objectives, and the Decision components, these will be described in this section.

Hazard Identification

As discussed in Section 4.2 the most essential hazards in a HRC is any form of unwanted contact between the human and the robot. This is therefore what will be regarded as the main hazard in this experiments, and the only one hazard to be included in the evaluation, that the human and the robot is occupying the same space at the same time.

Goals and Objectives

The primary objective was to keep the risk below a level of 0.5. Further, the system had a goal to complete as many tasks as possible, with an objective to not exceed an imbalance of one task being executed more than two times more than the others. The task progression is monitored and tasks with fewer executions was thus prioritized.

Decision

To investigate the contribution of the proactive safety layer, there are no other safety layers included in the system. Basic necessary safety measures such as emergency stop is included externally. The decision making mechanism in the system is thus only affected by the risk related objective and the productivity objective. The algorithm governing the decision component is depicted in Figure 3. Essentially, the system has four available options: execute Robot Task 1, Robot Task 2 or Robot Task 3, or it can wait. First, the tasks with a risk lower than the set threshold is identified. If no task has a low enough risk associated with it, the system waits. The number of executions of each task is then compared, and any task that has been executed two times or more than the least executed task is excluded from selection. The task with the combination of the lowest risk, and the lowest number of executions is then selected for execution. The system continuously runs this decision loop as long as the robot is in stand by. A system with all three safety layers would continuously run the decision loop, and decisions based on low level SA would overrule and interrupt actions started on the basis of higher level SA.

Algorithm 3 Decision Component

procedure SELECT ACTION
    if Robot busy = FALSE **then**
        if Risk(Task$_i$) <Risk$_{max}$ **then**
            if Task Penalty(Task$_i$) <Task Penalty$_{max}$ **then**
                Start Task$_i$
                Task penalty(Task$_i$) $\leftarrow$ Task Penalty(Task$_i$)+increment
            else Return
        else Return
    else Return

7.2.6 Parameter Settings

The workday was set to 10 minutes, and the maximum TBV set to 1 minute. Both the ToV and the TBV datasets were distributed over 200 bins, giving a time interval of 3 seconds and 0.3

seconds in (4.10) and (4.11) respectively. The maximum accepted MSE for the fitness of the curve fitting was set to $\varepsilon_{max} = 5 \times 10^{-4}$. The maximum number of PDFs was set to 5 and the relearn rate was set to 25%.

The design parameters used in the human motion predictions, k_1, k_2 and γ, were set to $k_1 = 1$, $k_2 = 0.25$ and $\gamma = 1$. The length of the human velocity prediction history, L, was set to $L = 1 second$. The maximum distance of the proximity field ρ was set to $\rho = 20cm$. The likelihood analysis iteration time was set to 1 second.

7.2.7 Participants

The experiments were conducted with a total of 9 participants. The participants were students at Chuo University and were all male in their 20's. The participants were recruited by email and word-of-mouth.

7.2.8 Hypothesis

A set of hypotheses was formulated for this experiment also. If the hypotheses hold, it is safe to conclude that the system corresponds to **PS3**. The effects on the human operator working with a Responsible Robot is tested regarding workload and stress. The workload is tested with the NASA-TLX, which was presented in Section 7.2.4. The NASA-TLX is a well used and reliable questionnaire working with several sub-scales that can be investigated separately as well as combined using a weighting scale. In these experiments, two set of weights will be used, one set by the participants, and one set by the research team. This is done to investigate both which sub-scales the participants find it the most important to improve, and to study the effects on the workload in the light of the research goals. The most important sub-scales to improve for the research team can be seen in Table 7.5.

Further, as a stress indicator, the number of human errors was counted in each case. Human errors are used as an indicator as they can be an indicator on lack of concentration. Causing many errors may also be frustrating and reduce the sense of accomplishment for the human. The errors included counting the wrong number of bricks, building the figure wrong, loosing bricks of the table, etc. Some smaller errors like missing the box with a brick when packing bricks or having to recount the bricks during picking, were counted as half an error. The goal of the system is thus to reduce the workload and the number of human errors and two hypotheses are formulated as shown below.

$H_{3.1}$ The proposed system will reduce the perceived workload for the human operator compared to a system using a preprogrammed robot co-worker.

$H_{3.2}$ The system reduces the number of human errors in the scenario.

7.3 Results

The experiments were conducted as previously described with a total of 9 participants. The HHC case was always conducted first, to give the system time to learn, then either the PP-HRC or Responsible Robots based-HRC was second. A total of 5 participants followed the HHC→PP-HRC→Responsible Robots based-HRC pattern, and 4 followed the HHC→ Responsible Robots based-HRC→PP-HRC pattern. Due to practical reasons the learning time allowed to the system was limited to a minimum, and the Responsible Robots based-HRC case was conducted after the first or second iteration, typically.

As mentioned, the workload and stress levels for the operators were studied through the NASA-TLX forms completed by the participants and the human error counts. The error counts were found by studying the video material. The errors were counted as either a full error or a half error. The two error types are previously described in Section 7.2.8, however some types of errors were on some occasions counted as a full error, and on some as a half. This was based on a subjective assessment based on the severity. In some cases did a recount of bricks only take a few seconds, while it sometimes took multiple seconds and caused more frustration for the human. The number of human errors for each of the participants in each case can be seen in Table 7.6.

Table 7.6: Number of human errors observed during the experiments.

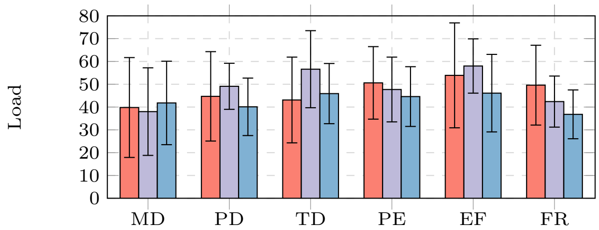
	HHC			PP-HRC			RR-HRC		
#	Full	Half	**Tot**	Full	Half	**Tot**	Full	Half	**Tot**
1	0	0	**0**	0	1	**0.5**	2	0	**2**
2	4	2	**5**	3	4	**5**	2	3	**3.5**
3	2	3	**3.5**	0	9	**4.5**	0	2	**1**
4	0	0	**0**	0	0	**0**	0	0	**0**
5	1	2	**2**	3	0	**3**	0	0	**0**
6	4	1	**4.5**	1	7	**4.5**	0	1	**0.5**
7	0	1	**0.5**	0	0	**0**	0	0	**0**
8	1	0	**1**	0	1	**0.5**	2	0	**2**
9	0	0	**0**	0	3	**1.5**	0	0	**0**

From examining the total errors the Responsible Robots based-HRC approach reduced the number of errors in 5 out of 9 experiments. Two of the 9 participants (4 and 7) had no errors in both the PP-HRC approach and the Responsible Robots based-HRC approach. Only participant 1 and 8 had more human errors with the Responsible Robots based-HRC approach compared to the PP-HRC approach. Another observation is the high number of human errors when the participant was working with another human in the HHC approach. This can most likely be accounted for by it being the first in the series. The average reduction or increase of numbers when comparing the approaches is shown in table 7.7. The mean reduction of errors with the Responsible Robots based-HRC approach compared to the two other approaches is close to one error. In the best experiment the number of errors was reduced with 3.5 errors when using the Responsible Robots based-HRC approach.

Table 7.7: Differences in number of human error between methods

	HHC - PP-HRC	HHC - RR-HRC	RR-HRC - PP-HRC
Mean	0.11	-0.83	-0.94
SEM	0.24	0.62	0.62
SD	0.74	1.85	1.86
Minimum	-0.5	-2	-3.5
Maximum	1.5	4	1.5
Count	9	9	9

Investigating the NASA-TLX forms gives insight in the experienced workload by the participants. The mean loads in each scale can be seen in Figure 7.3. The data reveals that the greatest improvement can be seen in temporal demand, effort and frustration. The mean of the weights defined by the participants are listed by priority in Table 7.8. Frustration is the highest weighted load by the participants, and physical demand the least. The effort, performance, temporal demand and mental demand all follow in between with similar weights. The weighted workloads are calculated using the weights defined by the participants and the ones defined by the research group as shown in (7.1) and (7.2).



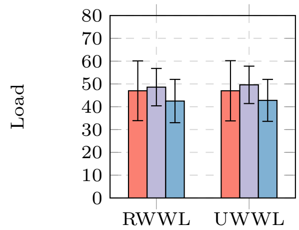
	MD	PD	TD	PE	EF	FR
HHC	39.8	44.7	43.1	50.6	53.9	49.6
PP-HRC	38.0	49.1	56.6	47.7	58.0	42.4
RAB-HRC	41.8	40.1	45.9	44.6	46.1	36.8

Figure 7.3: The mean of the scales investigated with the NASA-TLX

Table 7.8: Mean weight for the TLX scales

Weight Research	Weight User	Scale
5	5	Frustration
4	3.9	Effort
2	3.6	Performance
6	3.3	Temporal Demand
3	3.0	Mental Demand
1	2.2	Physical Demand

The means of the UWWL and RWWL can be seen in Figure 7.4. For both set of weights the UWWL and the RWWL are lower than both other approaches with the Responsible Robots based-HRC approach. It is also apparent that the two sets of weights produced similar sets of UWWLs and RWWLs.



		RWWL	UWWL
HHC		47	47
PP-HRC		48.6	49.6
RAB-HRC		42.5	42.8

Figure 7.4: The weighted means for the user and research weighted workloads

7.4 Evaluation of Results

The means of the UWWLs and RWWLs can be analyzed with a paired sample, two-tailed t-test. A comparison to the HHC approach is excluded since it was always tested first by the participants. This might have affected the level of perceived workload, since the participants then had little experience with the tasks.

The means of the UWWLs can be seen in Figure 7.4, comparing the means of PP-HRC and Responsible Robots based-HRC yeilds

$$\begin{aligned} t(9) &= 1.88, \\ p &= 0.039, \\ p < \alpha &\rightarrow \text{null hypothesis rejected}, \\ H_a : \Delta_{mean} \neq 0 &\rightarrow \text{is accepted.} \end{aligned}$$

Further, a comparison of the means of the RWWLs for PP-HRC and Responsible Robots based-HRC gives

$$\begin{aligned} t(9) &= 1.57, \\ p &= 0.068, \\ p < \alpha &\rightarrow \text{null hypothesis rejected}, \\ H_a : \Delta_{mean} \neq 0 &\rightarrow \text{is accepted.} \end{aligned}$$

As a result of the null hypothesis being rejected in both cases, it is safe to conclude that the Responsible Robots based-HRC approach present a lower workload for the operator than the PP-HRC approach and that hypothesis $\mathbf{H_{3.1}}$ holds.

The last indicator analyzed was the number of human errors throughout the workday. Similarly to the analysis of the number of conflicts the mean was compared to a mean of 0 in a one-tailed t-test. The t-test yielded

$$t(9) = 1.52,$$
$$p = 0.08,$$
$$p < \alpha \rightarrow \text{null hypothesis rejected},$$
$$H_a : \Delta_{mean} < 0 \rightarrow \text{is accepted},$$

and it is concluded that the Responsible Robots based-HRC approach reduces the number of human errors. Although the number of human errors is only one indicator it is a clear indication that hypothesis $\mathbf{H_{3.2}}$ holds. What caused these errors needs further investigations, however, when used as an indicator on concentration and stress the results support the use of the Responsible Robots based-HRC approach.

7.5 Summary

This experiment was conducted to investigate the proposed system's compliance with **PS3**.

PS3: *The developed system should be designed to improve the effect the collaboration has on the human operator.* The developed system should reduce the workload for the human and have a low rate of false alarms. Further, it should inherit some human-like attributes to build trust.

The effects of collaboration with a responsible robot compared to a pre-programmed robot have been investigated in this chapter. Two hypotheses were formulated to investigate these effects. Firstly, $\mathbf{H_{3.1}}$ stated that the system should reduce the perceived workload for the human operator. Further, $\mathbf{H_{3.2}}$ stated that the system would reduce the number of human errors due to a reduced stress level. Several participants worked in collaboration with a human, a pre-programmed robot and a responsible robot on an assembly task. The participants' errors were monitored in all cases, and task load scores were calculated.

Both the NASA-TLX workloads, UWWL and RWWL, were significantly reduced by the use of the proposed method. The UWWL was reduced from a load of 49.6 to 42.8, while the RWWL was reduced from 48.6 to 42.5. The reductions are a clear indication on the effects of working with a Responsible Robot and $\mathbf{H_{3.1}}$ was kept. The number of average errors was reduced by ~ 1. This might not seem like much, however, the experiment only lasted 10 minutes. Using the number of human errors as a stress indicator, the result is an indication of a reduced stress level. The hypothesis $\mathbf{H_{3.2}}$ is therefore kept.

The proposed approach proves to be more comfortable and less stressful to work with for humans than a preprogrammed robot and it is safe to conclude that **PS3** is fulfilled.

Chapter 8

Conclusion

8.1 Conclusion

In this work, a novel strategy for safe and productive human-robot collaboration has been presented, called Responsible Robots. The term Responsible Robots was chosen because rather than blindly obeying rules to ensure the safety of the human operator, the system makes decisions on the basis of which actions are safe.

A model was then proposed to realize Responsible Robots. The model enhances the robot's situation awareness by implementing a risk perception. The risk perception is based on standardized risk analysis framework and is active throughout operation. The enhanced situation awareness requires a projection of the future status of the system, which in this system is provided by the likelihood analysis in the risk perception. Three problem statements were formulated as criteria to the new system. The three problem statements regarded the proactive actions, productivity and the effect the proposed system had on the human operator.

The performance of the system was extensively tested in a series of experiments with human test subjects. The human test subjects received instructions that could be solved in a variety of ways. This variety represent the flexibility in how humans solve tasks and poses a tremendous challenge for the robotic system. Despite this, the experiments demonstrated that the system acts proactively against dangers with a precision of 96% and that the human operator's NASA-TLX workload is reduced by 14,5%. Moreover, the Responsible Robot reduced the number of human-robot conflicts by 81%.

It is therefore concluded that Responsible Robots as an approach to safe and productive HRC has been realized and that this approach has a positive effect on the human operator. The proposed method is also appropriate as a new layer of safety before the currently researched separation monitoring. The following sections will further discuss the conclusion about the problem statements and Responsible Robot-model. Lastly, some thoughts and ideas about future work is discussed in Section 8.2.

8.1.1 On Responsible Robots

A new strategy for safe and productive HRC called Responsible Robots has been proposed in this thesis. Responsible Robots have been introduced as robots that acts proactively against danger while maintaining productivity. A model to realize Responsible Robots was then presented and some of the possibilities that comes with the implemented risk perception was explored.

To realize a Responsible Robot some aspects about how human's make safe decisions has been explored. The importance of situation awareness in the human's decision making process was discussed. Further, the three levels of situation awareness have been presented and related to current research and available safety systems. It was found that a proper safety system at Level 3 SA was missing. The risk perception was identified as a means of enhancing the SA to Level 3 as the likelihood analysis give a projection of the future status of the system. The industrial standard risk framework was then presented and current safety systems was discussed in the light of the risk framework.

The relevant components in the Responsible Robot-model has been presented. The importance of the likelihood analysis was emphasized as it is the component that accounts for the projection of the future status of the system, as required at Level 3 SA. Models for realizing the Computation/Learning, Risk Estimate, Hazard Identification, Goals and Objectives, and the Decision component have been proposed.

8.1.2 On the Proactive Behavior of the Responsible Robot

PS1: *The developed system should act proactively against dangers.* Today's safety systems moves the robot away if it is in a conflict with the human to avoid a collision. The developed system should avoid human-robot conflicts, thus acting as a new layer of safety.

This PS was verified in two steps, first by purely testing the likelihood analysis, then through the decision maker in the completed system.

Two hypotheses were presented for the likelihood analysis, firstly that the system is able to predict whether the human operator will occupy a given SSW within the robots task execution time. Secondly, that the system's predictive capabilities would improve over time, as more observations were made, and the Computation/Learning component was given more time to calculate appropriate parameters. This was verified using evaluations of the DBS and BSS results and it was demonstrated that both hypotheses holds.

Furthermore, the proactive capabilities of the proposed Responsible Robots based HRC was experimentally tested with several test subjects. The participants worked alongside a responsible robot on an assembly task, and the robotic system decided autonomously when to safely execute its own tasks and when to wait. Two hypotheses were formulated to investigate this performance. Firstly, that the system had a low rate of false positives and secondly that the system reduced the number of human robot conflicts. It was shown through the experiment that both hypotheses holds.

On the basis of the two experiments and the validity of the four hypotheses it has been shown that the system is able to act proactive against dangers. It is concluded that the proposed system complies with **PS1**.

8.1.3 On the Productivity of the Responsible Robot

PS2: *The developed system should be able to solve the necessary tasks to maintain its productivity.* The system should be designed to be independent of task and robotic hardware. Further, the developed system should have an awareness of what the human operator expects of it.

The compliance of **PS2** was tested as a part of the experiment in Chapter 6. A HRC work cell was set up with a Responsible Robot and experimentally tested with several test subjects. The participants worked alongside a responsible robot on an assembly task, and the robotic system decided autonomously when to safely execute its own tasks and when to wait. A hypothesis was formulated that stated that the system is able to keep up with the human operator's productivity. An alpha level of 0.1 was arbitrarily chosen for evaluation purposes. The experiments demonstrated that also this hypothesis holds and it was verified that the Responsible Robot is able to keep up with the human's work pace. It is therefor safe to conclude that the proposed system is in compliance with **PS2**.

8.1.4 On the Effect of Working with a Responsible Robot

PS3: *The developed system should be designed to improve the effect the collaboration has on the human operator.* The developed system should reduce the workload for the human and have a low rate of false alarms. Further, it should inherit some human-like attributes to build trust.

The last PS stated that the system should have a positive effect on the human worker. This was tested in a similar experiment as the previously described scenarios. Firstly, two hypotheses were formulated to investigate this effects. It was stated that the system should reduce the perceived workload for the human operator and that the system would reduce the number of human errors. Several participants worked in a collaboration with a human, a pre-programmed robot and a Responsible Robot on an assembly task. The participant's errors were monitored in all cases, and task load scores were calculated using a NASA-TLX form. It was found that both hypotheses holds.

Therefore, the proposed approach proves to be more comfortable and less stressful to work with for the human operator than a preprogrammed robot. Therefore, it is safe to conclude that the proposed system complies with **PS3**.

8.2 Recommendations for further work

The system proposed in this thesis fulfilled all the problem statements defined in the beginning of the thesis. In spite of that, there are several approaches to improve the system. One of the great advantages of the system is its module based approach. The dynamic decision making model makes it possible to e.g. solely improve the Computations/Learning component. At some level, every component can always be improved both with regards to computational needs and performance.

More importantly, the strategy behind Responsible Robots can have an impact on research on HRC. Nowadays, most approaches focus either on safety or on task related challenges. The proposed strategy brings these two branches closer together with its safe task execution. This new way of thinking opens up for some new approaches to HRC in research with a greater focus on the entire collaboration.

Moreover, the proposed strategy is not dependent on or limited to industrial HRC and can be applied to any branch of robotics. One can even imagine Responsible Robots being implemented in any industrial robot. During installation of a new industrial robot, the Responsible Robot system would be aware of its lack of data about its environment and human activities. This way, it could be more prone to sound alarms when moving and alert the human integrator. The risk perception in general is a new modality on which human-robot communication can be based. Research into how the risk perception can be used to improve the robot's selectivity in communicated information would be interesting.

The Responsible Robots has been shown to have a positive effect on the human operator. However, a deeper investigation of the long-term effects of working with Responsible Robots would be beneficial. These investigations could include bio-data to get a better understanding of stress factors, concentration levels and the comfort of the human operator. After all, the human operator and the robot now the share the responsibility for the work cell's safety and productivity and should thus be devoted the same attention.

Acknowledgment

Many people have been involved in my PhD project and have contributed with insight, guidance and encouragement in the research process and the industrial aspects. I would like to thank these people for their support.

First of all, I would like to thank my supervisors Associate Professor Mihoko Niitsuma and Trygve Thomessen for invaluable guidance, commitment and inspiration throughout the project, we've had many important discussions. Many have asked my how I'm able to complete a degree in Japan without speaking Japanese and without the tremendous support of Niitsuma it surely would've been a lot more difficult. It has certainly been challenging at times, but very well worth it! Thomessen's encouragements have been very important, his mood and enthusiasm rubs of on everyone he meets.

I would also like to thank everyone at the Human-System Lab at Chuo University. Especially Keita and Yoshiaki for their support and unwavering translation of Japanese forms, documents and e-mails, and for their introduction to life in Japan and the amazing Japanese cuisine.

My colleagues at PPM also deserve a thanks for interesting discussions and insight. I would like to thank them for sharing their knowledge and experience, and for creating a great atmosphere at work.

My family has played an important role in encouraging and supporting me. Although I've been away for long periods, it only gave them a reason to visit Japan. Especially, Line deserves a special thanks for her support, patience and understanding.

This research was founded by The Research Council of Norway and PPM AS through the Industrial PhD program, which is gratefully acknowledged. I have also received financial support for my time in Japan from Hans Mustad og Robert og Ella Wenzins legat ved Norges Teknisk- naturvitenskapelige Universitet, for which I am grateful.

A.R.S

Bibliography

[1] H. W. Chesbrough, "The era of open innovation.," *MIT Sloan Management Review*, vol. 44, no. 3, pp. 35–41, 2003, ISSN: 15329194. DOI: 10.1371/journal.pone.0015090.

[2] A. R. Sanderud and T. Thomessen, "Releasing the synergy of human-robot collaboration - redundant robotics in practice," *ACTA Tehnica Corviniensis - Bulletin of Engineering*, vol. 7, no. 1, pp. 161–164, 2014.

[3] RIA/ANSI, *R15.06 1999 american standard for industrial robots safety requirement*, 1999.

[4] K. Behnisch, *White paper safe collaboration with abb robots electronic position switch and safemove*, 2008.

[5] *Robotics and robotic devices - safety requirements for industrial robots - part 1: robots*, 2011.

[6] *Robots and robotic devices - safety requirements for industrial robots - part 2: robot systems and integration*, 2011.

[7] *Robots and robotic devices – industrial safety requirements – collaborative industrial robots*, 2014.

[8] I. Asimov, *I, Robot*. Gnome Press, 1950, ISBN: 05-5390-033-1.

[9] E. Helms, R. D. Schraft, and M. Hägele, "Rod@work: robot assistant in industrial environments," in *Proceedings of the 2002 IEEE Int. Workshop on Robot and Human Interaction Communication*, 2002, pp. 399–404, ISBN: 0780375459.

[10] M. Hägele, W. Schaaf, and E. Helms, "Robot assistants at manual workplaces - effective co-operation and safety aspects," in *Proceedings of the 33rd International Symposium on Robotics (ISR)*, 2002, p. 6.

[11] C. Heyer, "Human-robot interaction and future industrial robotics applications," in *IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings*, Ieee, Oct. 2010, pp. 4749–4754, ISBN: 9781424466757. DOI: 10.1109/IROS.2010.5651294.

[12] J. Fryman and B. Matthias, "Safety of industrial robots: from conventional to collaborative applications," in *Robotics; Proceedings of ROBOTIK 2012; 7th German Conference on*, 2012, pp. 1–5, ISBN: 978-3-8007-3418-4.

[13] J. A. Marvel, J. Falco, and I. Marstio, "Characterizing task-based human – robot collaboration safety in manufacturing," *IEEE Transactions on systems, Man, and Cybernetics: Systems*, vol. 45, no. 2, pp. 260–275, 2015.

[14] C. W. Morato, K. N. Kaipa, J. Liu, and S. K. Gupta, "A framework for hybrid cells that support safe and efficient human-robot collaboration in assembly operations," in *Volume 1A: 34th Computers and Information in Engineering Conference*, ASME, Aug. 2014, V01AT02A078, ISBN: 978-0-7918-4628-5. DOI: 10.1115/DETC2014-34671.

[15] K. Eder, C. Harper, and U. Leonards, "Towards the safety of human-in-the-loop robotics: challenges and opportunities for safety assurance of robotic co-workers," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*, vol. abs/1404.2, Ieee, Aug. 2015, pp. 257–258, ISBN: 9781479967650. DOI: 10.1109/ROMAN.2014.6926328. arXiv: 1404.2229.

[16] R. Stiefelhagen, C. Fügen, P. Gieselmann, H. Holzapfel, K. Nickel, and A. Waibel, "Natural human-robot interaction using speech, head pose and gestures," in *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, vol. 3, 2004, pp. 2422–2427, ISBN: VO - 3.

[17] M. Rahman and Y. Wang, "Dynamic emotion-based human-robot collaborative assembly in manufacturing," in *Workshop on Human-Robot Collaboration for Industrial Manufacturing at RSS '14*, 2014.

[18] D. J. Bruemmer, D. a. Few, R. L. Boring, J. L. Marble, M. C. Walton, and C. W. Nielsen, "Shared understanding for collaborative control," *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, vol. 35, no. 4, pp. 494–504, Jul. 2005, ISSN: 10834427. DOI: 10.1109/TSMCA.2005.850599.

[19] D. a. Few, D. J. Bruemmer, and M. C. Walton, "Dynamic leadership for human-robot teams," in *Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction - HRI '06*, New York, New York, USA: ACM Press, 2006, p. 333, ISBN: 1595932941. DOI: 10.1145/1121241.1121302.

[20] J. R. Keebler, F. Jentsch, T. Fincannon, and I. Hudson, "Applying team heuristics to future human-robot systems," in *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on*, New York, New York, USA: ACM Press, 2012, pp. 169–170, ISBN: 2167-2121. DOI: 10.1145/2157689.2157739.

[21] I. Gris, D. A. Rivera, and D. Novick, "Animation guidelines for believable embodied conversational agent gestures," in *5th International Conference, VAMR 2013, Held as Part of HCI International 2013*, R. Shumaker, Ed., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, ch. Cybernetic, p. 399, ISBN: 978-3-642-39404-1. DOI: 10.1007/978-3-642-39405-8.

[22] S. Nikolaidis and J. Shah, "Human-robot cross-training: computational formulation, modeling and evaluation of a human team training strategy," in *ACM/IEEE International Conference on Human-Robot Interaction*, 2013, pp. 33–40, ISBN: 9781467330558. DOI: 10.1109/HRI.2013.6483499.

[23] M. A. Goodrich, T. W. McLain, J. D. Anderson, J. Sun, and J. W. Crandall, "Managing autonomy in robot teams: observations from four experiments," in *ACM/IEEE International Conference on Human-Robot Interaction*, 2007, pp. 25–32, ISBN: 9781595936172. DOI: 10.1145/1228716.1228721.

[24] J. Marble, D. Bruemmer, and D. Few, "Lessons learned from usability tests with a collaborative cognitive workspace for human-robot teams," in *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483)*, vol. 1, 2003, ISBN: 0-7803-7952-7. DOI: 10.1109/ICSMC.2003.1243856.

[25] J. L. Marble, D. J. Bruemmer, D. A. Few, and D. D. Dudenhoeffer, "Evaluation of supervisory vs. peer-peer interaction with human-robot teams," in *Proceedings of the Hawaii International Conference on System Sciences*, vol. 37, 2004, pp. 2067–2076, ISBN: 0-7695-2056-1. DOI: 10.1109/HICSS.2004.1265326.

[26] R. Hoffman, J. Lee, D. Woods, N. Shadbolt, J. Miller, and J. Bradshaw, "The dynamics of trust in cyberdomains," *IEEE Intelligent Systems*, vol. 24, no. 6, pp. 5–11, 2009, ISSN: 1541-1672. DOI: 10.1109/MIS.2009.124.

[27] R. Hoffman, M. Johnson, J. Bradshaw, and A. Underbrink, "Trust in automation," *IEEE Intelligent Systems*, vol. 28, no. 1, pp. 84–88, 2013, ISSN: 1541-1672. DOI: 10.1109/MIS.2013.24.

[28] R. C. Mayer, "An integrative model of organizational trust.," *The Academy of Management review*, vol. 20, no. 3, pp. 709–734, 1995.

[29] B. M. MUIR and N. MORAY, "Trust in automation. part ii. experimental studies of trust and human intervention in a process control simulation," *Ergonomics*, vol. 39, no. 3, pp. 429–460, 1996, ISSN: 0014-0139. DOI: 10.1080/00140139608964474.

[30] J. D. Lee and K. a. See, "Trust in automation: designing for appropriate reliance.," *Human factors*, vol. 46, no. 1, pp. 50–80, 2004, ISSN: 0018-7208. DOI: 10.1518/hfes.46.1.50.30392.

[31] A. H. Crusco and C. G. Wetzel, "The midas touch," *Personality and Social Psychology Bulletin*, vol. 10, no. 4, pp. 512 –517, Dec. 1984, ISSN: 0146-1672. DOI: 10.1177/0146167284104003.

[32] H. Fukuda, M. Shiomi, K. Nakagawa, and K. Ueda, "'midas touch' in human-robot interaction: evidence from event-related potentials during the ultimatum game," in *Proceedings of Human-Robot Interaction 2012*, vol. 2, 2012, pp. 131–132, ISBN: 9781450310635.

[33] E. Short, J. Hart, M. Vu, and B. Scassellati, "No fair!! an interaction with a cheating robot," in *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*, Ieee, Mar. 2010, pp. 219–226, ISBN: 978-1-4244-4892-0. DOI: 10.1109/HRI.2010.5453193.

[34] J. Fasola and M. J. Mataric, "Robot motivator: improving user performance on a physical/mental task," in *Human-Robot Interaction (HRI), 2009 4th ACM/IEEE International Conference on*, 2009, pp. 295–296, ISBN: 2167-2121. DOI: http://doi.acm.org/10.1145/1514095.1514183.

[35] M. Zinn, O. Khatib, B. Roth, and J. K. Salisbury, "Playing it safe," *IEEE Robotics & Automation Magazine*, vol. 11, no. 2, pp. 12–21, Sep. 2004, ISSN: 1070-9932. DOI: 10.1109/MRA.2004.1310938.

[36] J. Barraquand and J.-C. Latombe, "Robot motion planning: a distributed representation approach," *The International Journal of Robotics Research*, vol. 10, no. 6, pp. 628–649, 1991, ISSN: 0278-3649. DOI: 10.1177/027836499101000604.

[37] M. Stilman, "Task constrained motion planning in robot joint space," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Ieee, Oct. 2007, pp. 3074–3081, ISBN: 978-1-4244-0911-2. DOI: 10.1109/IROS.2007.4399305.

[38] Z. Yao and K Gupta, "Path planning with general end-effector constraints: using task space to guide configuration space search," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, Ieee, 2005, pp. 1875–1880, ISBN: 0-7803-8912-3. DOI: 10.1109/IROS.2005.1545305.

[39] D. Berenson, S. Srinivasa, and J. Kuffner, "Task space regions: a framework for pose-constrained manipulation planning," *The International Journal of Robotics Research*, vol. 30, no. 12, pp. 1435–1460, 2011, ISSN: 0278-3649. DOI: 10.1177/0278364910396389.

[40] J. Ahuactzin and K. Gupta, "The kinematic roadmap: a motion planning based global approach for inverse kinematics of redundant robots," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 4, pp. 653–669, 1999, ISSN: 1042-296X. DOI: 10.1109/70.781970.

[41] Z. Y. Guo and T. C. Hsia, "Joint trajectory generation for redundant robots in an environment with obstacles," *Journal of Robotic Systems*, vol. 10, no. 2, pp. 199–215, 1993, ISSN: 07412223. DOI: 10.1002/rob.4620100204.

[42] G. Oriolo, M. Ottavi, and M. Vendittelli, "Probabilistic motion planning for redundant robots along given end-effector paths," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and System*, vol. 2, 2002, 1657–1662 vol.2, ISBN: 0-7803-7398-7. DOI: 10.1109/IRDS.2002.1043993.

[43] J. Heinzmann and A. Zelinsky, "Quantitative safety guarantees for physical human-robot interaction," *The International Journal of Robotics Research*, vol. 22, no. 7-8, pp. 479–504, Jul. 2003, ISSN: 0278-3649. DOI: 10.1177/02783649030227004.

[44] S. Haddadin, A Albu-Schäffer, A. Albu-Schaffer, and A. Albu-sch, "Evaluation of collision detection and reaction for a human-friendly robot on biological tissues," in *6th IARP/IEEE-RAS/EURON Workshop on Technical Challenges for Dependable Robots in Human Environments*, 2008.

[45] S. Haddadin, A. Albu-Schäffer, A. De Luca, and G. Hirzinger, "Collision detection and reaction: a contribution to safe physical human-robot interaction," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, 2008, pp. 3356–3363, ISBN: 9781424420582. DOI: 10.1109/IROS.2008.4650764.

[46] M. Indri, S. Trapani, and I. Lazzero, "A general procedure for collision detection between an industrial robot and the environment," in *IEEE International Conference on Emerging Technologies and Factory Automation*, Luxembourg, 2015, ISBN: 9781467379298.

[47] O. Khatib and L. E. MAITRE, "Dynamic control of manipulators operating in a complex environment," in *Symposium on Theory and Practice of Robots and Manipulators*, 1980, pp. 267–282.

[48] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," in *Proceedings. 1985 IEEE International Conference on Robotics and Automation*, vol. 2, Udine, Italy, 1985, pp. 500–505. DOI: 10.1109/ROBOT.1985.1087247.

[49] ——, "A unified approach for motion and force control of robot manipulators: the operational space formulation," *IEEE Transactions on Robotics and Automation*, vol. 3, no. 1, pp. 43–53, 1987, ISSN: 0882-4967. DOI: 10.1109/JRA.1987.1087068.

[50] T. Tsuji, H. Akamatsu, and M. Kaneko, "Non-contact impedance control for redundant manipulators using visual information," in *Proceedings of International Conference on Robotics and Automation*, vol. 3, 1997, pp. 2571–2576, ISBN: 0-7803-3612-7. DOI: 10.1109/ROBOT.1997.619348.

[51] T. Tsuji and M. Kaneko, "Noncontact impedance control for redundant manipulators," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 29, no. 2, pp. 184–193, 1999, ISSN: 1083-4427. DOI: 10.1109/3468.747853.

[52] L. Balan and G. M. Bone, "Real-time 3d collision avoidance method for safe human and robot coexistence," in *IEEE International Conference on Intelligent Robots and Systems*, 2006, pp. 276–282, ISBN: 142440259X. DOI: 10.1109/IROS.2006.282068.

[53] D. Kulic, "Safety for human-robot interaction," Doctor of Philosophy, The University of British Colombia, 2006, pp. 1–171, ISBN: 9780494130117; 0494130113.

[54] D. Kulić and E. a. Croft, "Safe planning for human-robot interaction," *Journal of Robotic Systems*, vol. 22, no. 7, pp. 383–396, 2005, ISSN: 07412223. DOI: 10.1002/rob.20073.

[55] ——, "Real-time safety for human-robot interaction," *Robotics and Autonomous Systems*, vol. 54, no. 1, pp. 1–12, Jan. 2006, ISSN: 09218890. DOI: 10.1016/j.robot.2005.10.005.

[56] D Kulic and E. Croft, "Affective state estimation for human–robot interaction," *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 991–1000, 2007, ISSN: 1552-3098. DOI: 10.1109/TRO.2007.904899.

[57] Y. Zhang and J. Wang, "Obstacle avoidance for kinematically redundant manipulators using a dual neural network," *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 34, no. 1, pp. 752–759, 2004.

[58] T. Petrič and L. Žlajpah, "Smooth continuous transition between tasks on a kinematic control level: obstacle avoidance as a control problem," *Robotics and Autonomous Systems*, vol. 61, no. 9, pp. 948–959, Sep. 2013, ISSN: 09218890. DOI: 10.1016/j.robot.2013.04.019.

[59] B. Lacevic and P. Rocco, "Kinetostatic danger field - a novel safety assessment for human-robot interaction," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Ieee, Oct. 2010, pp. 2169–2174, ISBN: 978-1-4244-6674-0. DOI: 10.1109/IROS.2010.5649124.

[60] ——, "Safety-oriented control of robotic manipulators-a kinematic approach," in *IFAC World Congress*, 2011, pp. 11 508–11 513.

[61] B. Lacevic, P. Rocco, and A. M. Zanchettin, "Safety assessment and control of robotic manipulators using danger field," *IEEE Transactions on Robotics*, vol. 29, no. 5, pp. 1257–1270, Oct. 2013, ISSN: 1552-3098. DOI: 10.1109/TRO.2013.2271097.

[62] M. P. Polverini, A. M. Zanchettin, and P. Rocco, "Real-time collision avoidance in human-robot interaction based on kinetostatic safety field," English, in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, Sep. 2014, pp. 4136–4141, ISBN: 978-1-4799-6934-0. DOI: 10.1109/IROS.2014.6943145.

[63] S.-Y. Lo, C.-A. Cheng, and H.-P. Huang, “Virtual impedance control for safe human-robot interaction,” *Journal of Intelligent & Robotic Systems*, 2015, ISSN: 0921-0296. DOI: 10.1007/s10846-015-0250-y.

[64] B. Nemec, L. Zlajpah, and B. Nemec, “Kinematic control algorithms for on-line obstacle avoidance for redundant manipulators,” in *IEEE/RSJ Intl. Conference on Intelligent Robots and Systems*, vol. 2, 2002, pp. 1898–1903, ISBN: 0-7803-7398-7. DOI: 10.1109/IRDS.2002.1044033.

[65] K. Arras, J. Persson, N. Tomatis, and R. Siegwart, “Real-time obstacle avoidance for polygonal robots with a reduced dynamic window,” in *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, vol. 3, Ieee, 2002, pp. 3050–3055, ISBN: 0-7803-7272-7. DOI: 10.1109/ROBOT.2002.1013695.

[66] N. Pedrocchi, M. Malosio, and L. M. Tosatti, “Safe obstacle avoidance for industrial robot working without fences,” *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, pp. 3435–3440, Oct. 2009. DOI: 10.1109/IROS.2009.5353980.

[67] S. Mohammad Khansari-Zadeh and A. Billard, “A dynamical system approach to real-time obstacle avoidance,” *Autonomous Robots*, vol. 32, no. 4, pp. 433–454, 2012, ISSN: 09295593. DOI: 10.1007/s10514-012-9287-y.

[68] S. M. Khansari-zadeh and A. Billard, “Realtime avoidance of fast moving objects : a dynamical system-based approach,” in *Workshop on Robot Motion Planning: Online, Reactive, and in Real-time 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2012, 2012.

[69] N. Pedrocchi, F. Vicentini, M. Matteo, and L. Molinari, “Safe human-robot cooperation in an industrial environment,” *International Journal of Advanced Robotic Systems*, vol. 10, p. 1, 2013, ISSN: 1729-8806. DOI: 10.5772/53939.

[70] G. B. Avanzini, N. M. Ceriani, A. M. Zanchettin, P. Rocco, and L. Bascetta, “Safety control of industrial robots based on a distributed distance sensor,” *IEEE Transactions on Control Systems Technology*, vol. 22, no. 6, pp. 2127–2140, 2014. DOI: Doi10.1109/Tcst.2014.2300696.

[71] A. Zube, “Combined workspace monitoring and collision avoidance for mobile manipulators,” in *IEEE International Conference on Emerging Technologies and Factory Automation*, Luxembourg, 2015, ISBN: 9781467379298.

[72] A. R. Sanderud, T. Thomessen, H. Hashimoto, H. Osumi, and M. Niitsuma, “An approach to path planning and real-time redundancy control for human-robot collaboration,” in *2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, Ieee, Jul. 2014, pp. 1018–1023, ISBN: 978-1-4799-5736-1. DOI: 10.1109/AIM.2014.6878214.

[73] A. R. n. Sanderud, “Task programming of redundant industrial robots - a virtually extended nullspace formulation verified through obstacle avoidance.,” Master’s Thesis, Norwegian University of Science and Technology, 2012, p. 161.

[74] C. E. Madan, A. Kucukyilmaz, T. M. Sezgin, and C. Basdogan, “Recognition of haptic interaction patterns in dyadic joint object manipulation,” *IEEE Transactions on Haptics*, vol. 8, no. 1, pp. 54–66, 2015.

[75] H. S. Koppula and A. Saxena, "Anticipating human activities for reactive robotic response," in *IEEE International Conference on Intelligent Robots and Systems*, 2013, p. 2071, ISBN: 9781467363587. DOI: 10.1109/IROS.2013.6696634.

[76] E. Gribovskaya, A. Billard, and C. Bouzyd, "Human motion prediction in a human-robot joint task," *Citeseer*, 2008.

[77] E. Gribovskaya, A. Kheddar, and A. Billard, "Motion learning and adaptive impedance for robot control during physical interaction with humans," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2011, pp. 4326–4332, ISBN: 9781612843865. DOI: 10.1109/ICRA.2011.5980070.

[78] K. Kronander and A. Billard, "Online learning of varying stiffness through physical human-robot interaction," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2012, pp. 1842–1849, ISBN: 9781467314039. DOI: 10.1109/ICRA.2012.6224877.

[79] M. Khoramshahi, A. Shukla, and A. Billard, "Cognitive mechanism in synchronized motion: an internal predictive model for manual tracking control (special session)," in *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2014, pp. 765–771, ISBN: 978-1-4799-3840-7. DOI: 10.1109/SMC.2014.6974003.

[80] C. Lenz, S. Nair, M. Rickert, A. Knoll, W. Rösel, J. Gast, A. Bannat, and F. Wallhoff, "Joint-action for humans and industrial robots for assembly tasks," in *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN*, 2008, pp. 130–135, ISBN: 9781424422135. DOI: 10.1109/ROMAN.2008.4600655.

[81] D. Nakhaeinia, P. Laferriere, P. Payeur, and R. Laganiere, "Safe close-proximity and physical human-robot interaction using industrial robots," in *2015 12th Conference on Computer and Robot Vision*, 2015, pp. 237–244, ISBN: 978-1-4799-1986-4. DOI: 10.1109/CRV.2015.38.

[82] J. Mainprice and D. Berenson, "Human-robot collaborative manipulation planning using early prediction of human motion," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Ieee, Nov. 2013, pp. 299–306, ISBN: 978-1-4673-6358-7. DOI: 10.1109/IROS.2013.6696368.

[83] W. Y. Kwon and I. H. Suh, "A temporal bayesian network with application to design of a proactive robotic assistant," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2012, pp. 3685–3690, ISBN: 9781467314039. DOI: 10.1109/ICRA.2012.6224673.

[84] ——, "Planning of proactive behaviors for human–robot cooperative tasks under uncertainty," *Knowledge-Based Systems*, vol. 72, pp. 81–95, 2014, ISSN: 09507051. DOI: 10.1016/j.knosys.2014.08.021.

[85] R. Alami, R. Chatila, A Clodic, S. Fleury, M. Herrb, V. Montreuil, and E. A. Sisbot, "Towards human-aware cognitive robots," in *National Conference on Artificial Intelligence - AAAI*, 2006, p. 8, ISBN: 1577352858.

[86] E. A. Sisbot, K. F. Marin-Urias, R. Alami, and T. Siméon, "A human aware mobile robot motion planner," *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 874–883, 2007, ISSN: 15523098. DOI: 10.1109/TRO.2007.904911.

[87] E. A. Sisbot, L. F. Marin-Urias, X. Broquère, D. Sidobre, and R. Alami, "Synthesizing robot motions adapted to human presence: a planning and control framework for safe and socially acceptable robot motions," *International Journal of Social Robotics*, vol. 2, no. 3, pp. 329–343, 2010, ISSN: 18754791. DOI: 10.1007/s12369-010-0059-6.

[88] G. Hoffman, G. Hoffman, C. Breazeal, and C. Breazeal, "Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team," *ACM SIGCHI/SIGART Human-Robot Interaction*, pp. 1–8, 2007, ISSN: 2167-2121. DOI: http://doi.acm.org/10.1145/1228716.1228718.

[89] S. Nikolaidis, P. Lasota, G. Rossano, C. Martinez, T. Fuhlbrigge, and J. Shah, "Human-robot collaboration in manufacturing: quantitative evaluation of predictable, convergent joint action," in *IEEE ISR 2013*, 2013, pp. 1–6, ISBN: 978-1-4799-1173-8. DOI: 10.1109/ISR.2013.6695625.

[90] P. A. Lasota, G. F. Rossano, and J. A. Shah, "Toward safe close-proximity human-robot interaction with standard industrial robots," English, in *2014 IEEE International Conference on Automation Science and Engineering (CASE)*, IEEE, Aug. 2014, pp. 339–344, ISBN: 978-1-4799-5283-0. DOI: 10.1109/CoASE.2014.6899348.

[91] P. A. Lasota and J. a. Shah, "Analyzing the effects of human-aware motion planning on close-proximity human-robot collaboration," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 57, no. 1, pp. 21–33, 2015, ISSN: 0018-7208. DOI: 10.1177/0018720814565188.

[92] O. C. Schrempf, U. D. Hanebeck, A. J. Schmid, and H. Wörn, "A novel approach to proactive human-robot cooperation," English, in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, IEEE, 2005, pp. 555–560, ISBN: 0780392752. DOI: 10.1109/ROMAN.2005.1513838.

[93] E. Meisner, V. Isler, and J. Trinkle, "Controller design for human-robot interaction," *Autonomous Robots*, vol. 24, no. 2, pp. 123–134, 2007, ISSN: 0929-5593. DOI: 10.1007/s10514-007-9054-7.

[94] K. Madhava Krishna, R. Alami, and T. Simeon, "Safe proactive plans and their execution," *Robotics and Autonomous Systems*, vol. 54, no. 3, pp. 244–255, 2006, ISSN: 09218890. DOI: 10.1016/j.robot.2005.10.008.

[95] T. Kitade, S. Satake, T. Kanda, and M. Imai, "Understanding suitable locations for waiting," in *ACM/IEEE International Conference on Human-Robot Interaction*, Ieee, Mar. 2013, pp. 57–64, ISBN: 9781467330558. DOI: 10.1109/HRI.2013.6483502.

[96] S.-Y. Chung and H.-P. Huang, "Predictive navigation by understanding human motion patterns," *International Journal of Advanced Robotic Systems*, vol. 8, no. 1, p. 1, 2011, ISSN: 1729-8806. DOI: 10.5772/10529.

[97] B. Mutlu, F. Yamaoka, T. Kanda, H. Ishiguro, and N. Hagita, "Nonverbal leakage in robots: communication of intentions through seemingly unintentional behavior," in *Human-Robot Interaction (HRI), 2009 4th ACM/IEEE International Conference on*, vol. 2, 2009, pp. 69–76, ISBN: 9781605584041. DOI: 10.1145/1514095.1514110.

[98] P. Slovic, "Perception of risk," *Science*, vol. 236, no. 4799, pp. 280–285, 1987. DOI: DOI:10.2307/1698637.

[99] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 32–64, 1995, ISSN: 00187208. DOI: 10.1518/001872095779049543.

[100] Norsok, *Risiko- og beredskapsanalyse*, September. Standards Norway, 2001, ISBN: 9282695492.

[101] H.-I. Yang, Y.-W. Yun, and G.-J. Park, "Design of a pedestrian protection airbag system using experiments," in *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, Sep. 2015. DOI: 10.1177/0954407015603854.

[102] S. Choi, L. Engineering, J. Jang, L. Engineering, C. Oh, and L. Engineering, *Simulation-based methodology for evaluating the effectiveness of integrated pedestrian protection systems choi , jang , and oh*, 2015.

[103] T. Thomessen and M. Niitsuma, "Cognitive human-machine interface with multi- modal man-machine communication," in *4th IEEE International Conference on Cognitive Intocommunications*, 2013, pp. 873–876, ISBN: 9781479915460.

[104] P. Galambos, "Vibrotactile feedback for haptics and telemanipulation: survey, concept and experiment," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 41–65, 2012, ISSN: 17858860.

[105] Y. Muramatsu, M. Niitsuma, and T. Thomessen, "Building a cognitive model of tactile sensations based on vibrotactile stimuli," in *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)*, Ieee, Dec. 2013, pp. 149–154, ISBN: 978-1-4799-1546-0. DOI: 10.1109/CogInfoCom.2013.6719231.

[106] A. Sanderud, T. Thomessen, H. Osumi, and M. Niitsuma, "A proactive strategy for safe human-robot collaboration based on a simplified risk analysis," *Modeling, Identification and Control: A Norwegian Research Bulletin*, vol. 36, no. 1, pp. 11–21, 2015, ISSN: 0332-7353. DOI: 10.4173/mic.2015.1.2.

[107] S. Haddadin, A. Albu-Schäffer, M. Frommberger, J. Rossmann, and G. Hirzinger, "The "dlr crash report": towards a standard crash-testing protocol for robot safety - part i: results," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2009, pp. 272–279, ISBN: 9781424427895. DOI: 10.1109/ROBOT.2009.5152711.

[108] ——, "The "dlr crash report": towards a standard crash-testing protocol for robot safety - part i: results," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2009, pp. 272–279, ISBN: 9781424427895. DOI: 10.1109/ROBOT.2009.5152711.

[109] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963, ISSN: 0368-4245. DOI: 10.1137/0111030. arXiv: arXiv:1011.1669v3.

[110] J. Versace, "A review of the severity index," in *In Proceedings of the 15th Stapp Car Crash Conference, SAE Paper No. 710881*, 1971, pp. 771–796. DOI: 10.4271/710881.

[111] T. Lobdell, *Human Impact Response*, W. F. King and H. J. Mertz, Eds. Boston, MA: Springer US, 1973, ISBN: 978-1-4757-1504-0. DOI: 10.1007/978-1-4757-1502-6.

[112] E Hertz, "A note on the head injury criterion (hic) as a predictor of the risk of skull fracture," in *Proceedings: Association for the Advancement of Automotive Medicine Annual Conference*, vol. 37, Plaines, IL, 1993, pp. 303–312, ISBN: 1540-0344.

[113] P. Prasad and H. J. Mertz, *The position of the united states delegation to the iso working group 6 on the use of hic in the automotive environment*, English, Jun. 1985. DOI: 10.4271/851246.

[114] *Ais for the advancement of automative medicine, the abreviated injury scale (1990)*, Plaines, IL, 1998.

[115] A. H. Murphy, "A new vector partition of the probability score," *Journal of Applied Meteorology*, vol. 12, no. 4, pp. 595–600, 1973, ISSN: 0021-8952. DOI: 10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2.

[116] G. W. Brier, "Verification of forecasts expersses in terms of probaility.," *Monthly Weather Review*, vol. 78, no. 1, pp. 1–3, 1950, ISSN: 0036-8075. DOI: 10.1126/science.27.693.594.

[117] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): results of empirical and theoretical research," *Advances in Psychology*, vol. 52, no. C, pp. 139–183, 1988, ISSN: 01664115. DOI: 10.1016/S0166-4115(08)62386-9.

Author's Publications

Journal Papers

[a] A. Sanderud and T. Thomessen, "Releasing the Synergy of Human-Robot Collaboration - Redundant Robotics in Practice." *ACTA Tehnica Corviniensis - Bulletin of Engineering*, 2014. Volume 7, Issue 1, pages 161-164

[b] A. Sanderud, T. Thomessen, H. Osumi and M. Niitsuma, "A Proactive strategy for safe human-robot collaboration based on a simplified risk analysis." *Modeling, Identification and Control: A Norwegian Research Bulletin*, vol. 36, no. 1, pp. 11-21, 2015, ISSN: 0332-7353. DOI: 10.4173/mic.2015.1.2.

[c] A. Sanderud, T. Thomessen and M. Niitsuma, "Likelihood Analysis for a Proactive Risk Analysis based safe Human-Robot Collaboration" *(In Review)*

[d] A. Sanderud, T. Thomessen and M. Niitsuma, "Human-Robot Collaboration with Responsible Robots" *(In Review)*

Conference Papers

[e] A. Sanderud, T. Thomessen, H. Hashimoto, H. Osumi and M. Niitsuma, "An approach to path planning and real-time redundancy control for human-robot collaboration." *Proceedings of 2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics* (AIM2014), pages 1018-1023.

[f] A. Sanderud, T. Thomessen and M. Niitsuma, "A Likelihood Analysis for a Risk Analysis for Safe Human Robot Collaboration" *Proceedings of 20th IEEE International conference on Emerging Technologies and Factory Automation*, 2015 (ETFA2015)

[g] A. Sanderud, T. Thomessen and M. Niitsuma, "Proactive Safety System Using Risk Analysis in a Human Robot Collaboration" *2nd European TA Conference on The Next Horizon of Technology Assessment*, 2015 (PACITA2015) *(without review)*

List of Figures

List of Tables

Appendix A

Results from Skill Testing in Chapter 5

Table A.1: Decomposed Brier Score (BS) and Brier Skill Scores (BSS) for Scenario 1 in Chapter 5

			SSW1		
Day	REL	RES	UNC	DBS	BSS
1	0.0519	0.0777	0.1875	0.1617	0.1379
2	0.0330	0.0808	0.1956	0.1478	0.2441
3	0.0259	0.0777	0.1908	0.1390	0.2713
4	0.0363	0.0758	0.1971	0.1576	0.2003
5	0.0391	0.0763	0.1940	0.1568	0.1916
6	0.0338	0.0767	0.2059	0.1629	0.2086
7	0.0257	0.0720	0.2001	0.1538	0.2314
8	0.0175	0.0761	0.1986	0.1400	0.2950
9	0.0142	0.0810	0.1986	0.1318	0.3363
10	0.0114	0.0844	0.2052	0.1323	0.3554
Avg.	0.0289	0.0778	0.1973	0.1484	0.2472

			SSW2		
Day	REL	RES	UNC	DBS	BSS
1	0.0055	0.1236	0.2495	0.1314	0.4736
2	0.0051	0.1505	0.2488	0.1034	0.5844
3	0.0038	0.1489	0.2495	0.1044	0.5815
4	0.0039	0.1455	0.2478	0.1062	0.5716
5	0.0046	0.1459	0.2487	0.1073	0.5684
6	0.0071	0.1361	0.2458	0.1168	0.5249
7	0.0033	0.1399	0.2475	0.1109	0.5519
8	0.0055	0.1630	0.2484	0.0910	0.6338
9	0.0051	0.1725	0.2485	0.0812	0.6734
10	0.0054	0.1391	0.2498	0.1161	0.5351
Avg.	0.0049	0.1465	0.2484	0.1069	0.5698

			SSW3		
Day	REL	RES	UNC	DBS	BSS
1	0.0055	0.1236	0.2488	0.1307	0.4747
2	0.0071	0.1557	0.2496	0.1009	0.5956
3	0.0068	0.1530	0.2491	0.1029	0.5868
4	0.0066	0.1523	0.2499	0.1043	0.5828
5	0.0055	0.1519	0.2497	0.1033	0.5864
6	0.0073	0.1420	0.2499	0.1152	0.5391
7	0.0068	0.1488	0.2500	0.1081	0.5677
8	0.0076	0.1658	0.2498	0.0916	0.6332
9	0.0076	0.1760	0.2498	0.0813	0.6744
10	0.0059	0.1366	0.2453	0.1146	0.5330
Avg.	0.0067	0.1506	0.2492	0.1053	0.5774

Table A.2: Decomposed Brier Score (BS) and Brier Skill Scores (BSS) for Scenario 2 in Chapter 5

			SSW1		
Day	REL	RES	UNC	DBS	BSS
1	0.0085	0.1840	0.2488	0.0732	0.7056
2	0.0115	0.2174	0.2500	0.0441	0.8237
3	0.0058	0.1729	0.2491	0.0820	0.6708
4	0.0153	0.2232	0.2495	0.0415	0.8337
5	0.0077	0.1870	0.2490	0.0697	0.7202
6	0.0094	0.1909	0.2495	0.0681	0.7271
7	0.0121	0.1979	0.2478	0.0620	0.7498
8	0.0104	0.2133	0.2498	0.0469	0.8124
9	0.0096	0.1985	0.2483	0.0594	0.7608
10	0.0095	0.1915	0.2481	0.0661	0.7338
Avg.	0.0100	0.1977	0.2490	0.0613	0.7538

			SSW2		
Day	REL	RES	UNC	DBS	BSS
1	0.0161	0.0489	0.2446	0.2118	0.1341
2	0.0117	0.0865	0.2477	0.1728	0.3022
3	0.0302	0.0383	0.2419	0.2338	0.0333
4	0.0181	0.1052	0.2478	0.1607	0.3514
5	0.0106	0.0663	0.2495	0.1938	0.2233
6	0.0221	0.0365	0.2475	0.2330	0.0584
7	0.0038	0.1165	0.2498	0.1371	0.4510
8	0.0072	0.1285	0.2499	0.1287	0.4850
9	0.0062	0.1116	0.2497	0.1442	0.4222
10	0.0036	0.0911	0.2500	0.1624	0.3504
Avg.	0.0130	0.0830	0.2478	0.1779	0.2811

			SSW3		
Day	REL	RES	UNC	DBS	BSS
1	0.0053	0.1621	0.2431	0.0862	0.6453
2	0.0054	0.1788	0.2449	0.0714	0.7083
3	0.0060	0.1513	0.2390	0.0937	0.6078
4	0.0126	0.2028	0.2456	0.0553	0.7746
5	0.0042	0.1488	0.2433	0.0988	0.5941
6	0.0040	0.1508	0.2449	0.0980	0.5996
7	0.0032	0.1733	0.2458	0.0757	0.6921
8	0.0085	0.1913	0.2446	0.0618	0.7472
9	0.0050	0.1726	0.2446	0.0771	0.6849
10	0.0045	0.1736	0.2468	0.0778	0.6848
Avg.	0.0059	0.1705	0.2442	0.0796	0.6739

Table A.3: Decomposed Brier Score (BS) and Brier Skill Scores (BSS) for Scenario 3 in Chapter 5

Day	SSW1 REL	RES	UNC	DBS	BSS
1	0.0268	0.0760	0.2480	0.1987	0.1986
2	0.0128	0.1040	0.2425	0.1513	0.3760
3	0.0134	0.1323	0.2413	0.1224	0.4927
4	0.0195	0.1111	0.2462	0.1546	0.3720
5	0.0153	0.1418	0.2410	0.1145	0.5248
6	0.0195	0.1657	0.2403	0.0941	0.6083
7	0.0192	0.1252	0.2484	0.1425	0.4265
8	0.0179	0.1584	0.2425	0.1019	0.5797
9	0.0147	0.1242	0.2475	0.1380	0.4425
10	0.0150	0.1009	0.2472	0.1613	0.3474
Avg.	0.0174	0.1240	0.2445	0.1379	0.4368

Day	SSW2 REL	RES	UNC	DBS	BSS
1	0.0173	0.1693	0.2497	0.0978	0.6085
2	0.0184	0.1825	0.2495	0.0854	0.6577
3	0.0184	0.1786	0.2495	0.0892	0.6423
4	0.0176	0.1682	0.2496	0.0990	0.6034
5	0.0243	0.1525	0.2460	0.1178	0.5211
6	0.0188	0.1784	0.2496	0.0900	0.6394
7	0.0110	0.1748	0.2495	0.0857	0.6564
8	0.0107	0.1752	0.2499	0.0854	0.6582
9	0.0145	0.1844	0.2495	0.0796	0.6811
10	0.0150	0.1770	0.2487	0.0866	0.6516
Avg.	0.0166	0.1741	0.2491	0.0917	0.6320

Day	SSW3 REL	RES	UNC	DBS	BSS
1	0.0101	0.1480	0.2473	0.1095	0.5574
2	0.0082	0.1626	0.2480	0.0936	0.6224
3	0.0122	0.1838	0.2484	0.0768	0.6908
4	0.0161	0.1677	0.2481	0.0965	0.6109
5	0.0199	0.1944	0.2483	0.0737	0.7031
6	0.0190	0.1906	0.2480	0.0763	0.6921
7	0.0159	0.1871	0.2478	0.0766	0.6908
8	0.0130	0.1832	0.2464	0.0762	0.6906
9	0.0198	0.1924	0.2478	0.0752	0.6966
10	0.0178	0.1864	0.2490	0.0804	0.6773
Avg.	0.0152	0.1796	0.2479	0.0835	0.6632

Appendix B

Hardware and Software

The following hardware and software was used in the experimental setup of the proposed system.

Hardware

Sensor

Microsoft Kinect v2 for Windows, time of flight sensor. (dev.windows.com/kinect)

Robot

NACHI MR20 7-axes industrial robot. (www.nachirobotics.com/mr20.html)
NACHI FD11 robot controller. (www.nachirobotics.com/sales-data/fd-controller.html)
In-house high-speed interface to the FD11 controller.

Computers

Main Computer: Intel Core i7-950 @ 3.07GHz, 16GB RAM, Windows 8.1 Pro.
Computations/Learning-component Computer: Intel Core i7-5820K @ 3.30GHz, 16GB RAM, Windows 8.1 Pro.

Software

National Instruments LabVIEW 2012 SP1 Development System with, Real-Time Module, Robotics Module, Vision Development Module, MathScript RT Module. (www.ni.com/labview)
HARO 3D library for Kinect v2. (www.harotek.com)