A Novel Decision-Making Architecture for Human-Robot Collaboration

From Mirror Neurons to Conflict-free Interactions

By

MEHDI SOBHANI

School of Engineering

COLLEGE OF ARTS, TECHNOLOGY AND ENVIRONMENT

UNIVERSITY OF THE WEST OF ENGLAND

A PhD dissertation submitted to the University of the West of England in accordance with the requirements of the degree of DOCTOR IN PHILOSOPHY in the College of Arts, Technology and Environment.

SEPTEMBER 2023

Abstract

Inspired by the role of mirror neurons and the importance of predictions in joint action, a novel decision-making structure is proposed, designed and tested for both individual and dyadic action during real-world human-human and human-robot experiments. The structure comprises models representing individual decision policies, policy integration layer(s), and a negotiation layer. The latter is introduced to prevent and resolve conflicts among individuals through internal simulation rather than via explicit agent-agent communication.

As the main modelling tool, Dynamic Neural Fields (DNFs) were chosen. Data was captured from human-human experiments with a decision-making task performed by either one or two participants. The task involves choosing (picking) and placing blocks one by one from seven wooden blocks to create an alpha/numeric character on a kind of mechanical model of a 7-segment display. The task is designed to be as generic as possible. Recorded hand and blocks movements were used for developing DNF-based models by optimising parameters using a genetic algorithm.

Results show that decision policies can be modelled and integrated with acceptable accuracy for individual performances. In the dyadic experiment, using only individual models without the negotiation layer, the model failed to resolve conflicts. However, with the implementation of a negotiation layer, this problem could be overcome.

To Analyse the proposed model for a human-robot collaboration task, first, the role and efficacy of the negotiation layer of the architecture are assessed. Then, in a "Wizard of Oz" experiment, the performance of the complete architecture is compared with that of a human decision-maker. The same task of using wooden blocks to create characters in a 7-segment display is used in both experiments.

Results show a significant improvement in terms of the chosen objective and subjective measures when the robot uses the complete architecture with the negotiation layer. No significant difference was found for any of the measures between the human decision-maker and the complete model. Although the robot with a human decision-maker scored slightly better in all measures, a further Bayesian comparison of the data suggests a high probability of similarity between the model and the human decision-maker. This was further illustrated by a qualitative analysis of the post-experiment interview questions; in answering the third question, when asked which condition is more human-like, 17 participants identified that the robot using the complete model was like working with a human, and an equal number opted for identifying the robot controlled by a human decision-maker as being human-like. In addition, answering the first question, 6 participants found no difference between the robot being controlled by a human decision-maker and being controlled using the complete model.

The proposed decision-making structure based on DNFs is developed and tested for a simple pick-and-place task. However, the main primitive underlying action of this task, pick-and-place, is indeed part of many more complex tasks people perform in their day-to-day life. Paired with the possibility to gradually evolve the architecture by adding new policies on demand, the architecture provides a general framework for modelling decision-making in joint action tasks.

To demonstrate the generality of this ability, a car assembling task was used in a "Wizard of Oz" experiment. Similar to the previous experiment, participants worked with a robotic arm to perform the task. Each participant repeated the task 6 times, 3 times for each condition, Model or Wizard, in a random order. Again, no significant difference was found between the two conditions and the Bayesian comparison showed a high probability of similarity. When data were sorted based on the order of trials, a significant difference was found between the task completion time from the first trial to the last. This could be due to the fact that participants were repeating the same task, however, given the low number of participants for this experiment, which was executed on a small scale only to illustrate the potential for the ability to transfer the capability to a different task, further analysis is required in the future.

List of publications

1. **Sobhani M.**, Smith J., Pipe A. and Peer A., (Accepted) **"A Decision-Making Architecture for Human-Robot Interaction: Model Transferability."** ICINCO 2023, 20th International Conference on Informatics in Control, Automation and Robotics. (The paper was double-blind peer reviewed by 4 reviewers)

2. **Sobhani M.**, Giuliani M., Smith J., Pipe A. and Peer A., (Accepted) **"Evaluating a Decision-Making Architecture in Human-Robot Collaboration Experiments."** International Journal of Social Robotics.

3. **Sobhani M.**, Smith J., Pipe A., Peer A., (2023) **"A Novel Mirror Neuron Inspired Decision-Making Architecture for Human-Robot Interaction."** International Journal of Social Robotics.

DEDICATION AND ACKNOWLEDGEMENTS

To the great people of Iran, a nation comprised of several ethnicities from Azeri, Baloch, Kurd, Lur, Persian, Turkmen and many more yet united in their ongoing struggle for freedom, democracy, and social justice since 1905 Constitution Revolution.

I would like to thank my supervisory team for their wisdom, guidance and support throughout my research project. I am indeed grateful to all the people helping me in the Bristol Robotics Laboratory and creating an amazing experience during a challenging period of my life. And my biggest thank to my family for their unconditional love and support.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE: ....

Table of Contents

Page

List of Tables

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

As they emerge from being dangerous caged tools, robots are becoming part of human day-to-day life. This has increased the direct interaction of humans and robots in different contexts from manufacturing to assisted living. It is well-researched that to have a successful and efficient human-robot collaboration (HRC) robots require to have several different cognitive abilities like perspective-taking [115], understanding affordances [82] (including situation, partner, and self-awareness), forming an expectation of the next action [79], and timing ability [25, 132]. These form a social cognitive process that starts at the perception level in which the robot assesses the situation and, by constant monitoring of the environment and partner(s), forms some understanding and predictions of the next action.

Different criteria have been suggested for the evaluation of a cognitive system like generality, versatility, rationality, optimality, efficiency, scalability, reactivity, persistence, improvability, autonomy, and extended operation [76]. Building on these criteria, several cognitive architectures in the literature, like ACT-R [4], Soar [73], or R-CAST, which are based on Recognition Primed Decision (RPD) models [42] are developed for individual agents. These architectures are either based on declarative memory retrieval using instance-based models or rule-based (ACT-R) or probabilistic modelling approaches like decision trees (Soar). One of the most recent applications of the ACT-R architecture in modelling decision-making is presented by Zhang et al. [135] for Human-Computer Interaction (HCI). While they developed a dynamic model for "complex" interactions in HCI, their model only produces a prediction of the individual's decision, but the missing embodied nature of robots and the collaborative nature of the task with a shared plan make such a model not necessarily applicable to HRC and joint-action scenarios. On the other hand, Wolpert et al. [128] developed a structure for action production or action observation of individuals. Although they argue that this concept of forward-inverse models used for modelling single motor control actions can be also applied to social interaction, their architecture has never been implemented for a real interaction scenario involving joint action.

One key part of the cognitive process is decision-making. According to Wang and Ruhe [123], the "decision criteria depend on decision strategy". They categorised decision strategies into four groups namely, intuitive, empirical, heuristic, and rational; each of which divided into different subsets. Decision makers may select different decision strategies even in the same circumstances depending on their "values", "attitude towards risk" and "prediction of the future outcome". Decision strategies are also divided into static or dynamic strategies. It is considered static when the environmental changes do not affect the decision-maker's action. On the other hand, when the decision maker's actions are influenced by environmental changes and vice versa, the decision strategy should be a dynamic one. For human-robot collaboration (HRC), for example, the strategies must be dynamic as each partner should consider the other partner's action and the progress of the shared plan when deciding about its action. In such a situation, decision-makers need information about the "actions", "the intention of peers", "the abilities of all partners" and "the state of the environment". In this work, only two of these namely, "actions" and " "the state of the environment" is used and "the intention of peers" is considered already known by agents. This information then can be used to choose the right action among all the possible actions [8]. These strategies can be utilised to implement a decision-making system based on different paradigms like the Game theory or the Bayesian theory [123]. However, as Vinciarelli et al. [119] point out "mutual influences" in the interaction process has not been well investigated.

Bicho et al. [16] proposed a decision-making system for joint action based on DNFs, however, in their work decision policies were hard-coded rather than being modelled from human experimental data. The workspace was divided into two sides to be able to predict an action to be performed by a co-actor. They assume that objects in the area closer to each actor (human or robot) will only be picked by the nearest actor.

In contrast, in developing the proposed architecture, the workspace is considered to be shared equally as it was observed in the human-human collaboration experiments that people did not necessarily act based on the assumption of a divided workspace and reached into their partner's area for picking objects. Furthermore, their decision-making system was only tested in joint action scenarios that involve serial actions with collaborators taking turns and performing complementary actions. This reduces potential conflicts significantly, while the proposed architecture is developed based on both serial and parallel actions with a negotiation layer to resolve conflicts. This means, if there is no physical constraint or limitations imposed by the shared plan, the actor can perform an independent action in parallel to his/her/its partner.

Related is also the recent work by Beraldo et al. [13] in the sense that they developed a decision-making structure by breaking down the process into decision policies. But the work lacks the joint action aspect of the interaction and has been developed for the teleoperation of a mobile robot. Finally, Buisan et al. [20] present an architecture for human-aware planning that emulates the partner's decision to better predict his/her actions as also done in the presented architecture in the next chapter. But both works [13, 20] miss a module for resolving conflicts in decision-making. Finally, the idea of mirror neurons has also been explored in literature when developing robot control architectures. Metta et al. [80], for example, developed a model of mirror neurons for robots that are supposed to learn to grasp from human observations.

This work is inspired by neuroscientific findings suggesting that an agent runs internal simulations whenever s/he attempts to perform an action or whenever an action is observed while being performed by someone else [128]. Since the 1980s the Simulation theory (ST), first presented by Gordon [53], along with other approaches like Theory theory (TT) and Rationality theory were competing to explain different aspects of human cognition. TT argues that people form a theory or an abstract model about their partner's mental states based on their experience, Rationality theory uses rationality principles to achieve this, whereas ST suggests that people internally simulate their partner's mental state to reach a "pretend" state (involving imaginative and empathetic processes as well as mirroring observed actions, for instance seeing an object being picked up the motor neurons of the observer are activated as if s/he is performing the picking up action.)[105]. Simulation theory has gained additional support in explaining cognitive processes of human interaction after the discovery of mirror neurons [10, 47]. ST has also inspired roboticists to develop cognitive architectures for safer [127] and more ethical [118] robots.

Further application of simulation theory of mind in robotics has also been reviewed by Bianco and Ognibene [15] as an approach to achieve "coordinating and managing false beliefs", "proactivity and preparation", "perception", and "learning" in social robots. Devin and Alami [33] also used ST for shared plan execution having the robot engage in a dialogue with its partner for a "fluent" implementation of the shared plan "without being annoying or intrusive" by estimating its partner's mental state. In addition, Görür et al. [55] propose an ST-based architecture for human intention estimation for robot decision-making in a shared plan. However, their work is presented in an online archive (not peer-reviewed) and the extent of the paper does not include any test either in the real world or simulation for assessing and evaluating the proposed architecture. Building on this work, they have presented an architecture for "anticipatory decision-making" [54] used to address two conditions: "1) when the human's intention is estimated to be irrelevant to the assigned task and may be unknown to the robot, e.g., motivation is lost, another assignment is received, onset of tiredness, and 2) when the human's intention is relevant but the human doesn't want the robot's assistance".

Particularly in the pre-motor cortex, two types of mirror neurons and canonical neurons have been found activated during action execution, imitation, or when only observing the other agent's actions. The mirror neurons were found to be activated during an action execution or observation with a specific goal, while the canonical neurons were found to be activated with the presentation of objects that afford goal-oriented actions [64, 117]. Inspired by the role of mirror neurons in joint action [90] and the fact that prediction is an essential part of this process, a novel decision-making architecture is proposed. Considering the importance of prediction in the joint action process in which one's own action system is used to understand and interact with others [10, 104] to enable an agent to form expectations about the next action of a collaborating partner, the proposed architecture foresees mental models of the decision-making processes of both the agent and of the interaction partner. Each agent has its own decision-making system that allows for combining a series of independent individual policies through an integration layer. The two decision-making systems of the agents run in parallel when collaborating on a joint action task and their outcome enters a negotiation layer. This layer is introduced to prevent conflicts in action execution by negotiating own independently taken decisions with anticipated partner's decisions. The latter is obtained by internal simulation of the mental model of the partner. So, each agent is assumed to simulate its own and its partner's decision-making process and to integrate the two independent decisions deriving from these processes into the final decision. The architecture is explained in detail in Chapter 3. A literature review is presented in the next chapter.

1.1 Research Hypotheses

As reviewed in the literature in the next chapter, to the best of our knowledge, there is no architecture for addressing decision-making challenges in dynamic human-robot collaboration. Therefore, in this work, a novel architecture is designed aiming to improve the collaboration by masking the interaction as conflict-free as possible. The conflict resolution is done through the implementation of the Negotiation layer which acts as an implicit communication. It is important to note there are some previous works [24, 28] in which conflict resolution was done through verbal communication. However, having verbal communication in HRC is not necessarily possible for many applications like working in a highly noisy environment or security patrolling. In addition, human beings rely on implicit communication along with verbal communication and this makes human-human collaboration fluent and seamless [50]. This makes the presented architecture in this work a novel and crucial development in advancing HRC.

Focusing on HRC and a joint action task, modelling decision-making policies is the first step in developing the architecture. After comparing different mathematical frameworks, Dynamic Neural Field (DNF) is chosen for the modelling policies and developing a decision-making architecture. As Curioni et al. [31] pointed out, human-human interaction can provide a good model for human-robot interaction (HRI), hence, a human-human collaboration experiment is designed and performed to capture the required data for the modelling. The experiment and modelling are done to investigate the following hypothesis:

- The decision process in dyadic joint action in a pick-and-place task can be broken down into decision policies and modelled accordingly.
- The decision-making models for the individuals can be integrated into a decision-making architecture for a dyadic joint action by adding a negotiation layer to resolve conflicts in a generic pick-and-place task.

- The models and architecture developed based on a generic and abstract pick-and-place task could be applied to a similar yet more realistic task without retraining the models.

After achieving a high level of accuracy in the modelling phase, the architecture is embedded in a robot for a human-robot experiment. The experiment is run in two phases, first to assess the effectiveness of the negotiation layer and then to compare the architecture to a human decision-maker. The result of the experiment shows that adding the negotiation layer has a significant effect in improving HRI and the robot using the architecture has a similar performance to the robot with a human decision-maker.

Finally, to show the transferability of the architecture, another experiment is designed to use the same models adapted for a more complex task without extra training. The result of this experiment shows that the architecture could be deployed for different tasks with small adjustments and no retraining.

After reviewing the literature in the next chapter, the architecture and the comparison of the potential mathematical frameworks for the decision-making models are presented in Chapter 3. The human-human experiment, data collection process, and developing models using DNFs are presented in Chapter 4. Two human-robot experiments are presented in Chapter 5. Chapter 6 presents the HRC experiment designed to show the generalisability and expandability of the architecture. Finally, the thesis is concluded and future work is presented in Chapter 7.

Chapter 2

Literature Review

As mentioned in the previous chapter, there are several cognitive processes required for a seamless human-robot collaboration. In this chapter, previous efforts in line with the proposed architecture are reviewed. First, related work on Action and Intention recognition is presented. Then, works related to physical interactions, like carrying an object, are reviewed and finally, the main cognitive architectures are presented and compared to the proposed architecture.

2.1 Action and Intention Recognition

As mentioned before, in a collaboration a decision maker needs to recognise his/her partner's actions and intentions ([8]). Intention recognition could be the next step after action recognition. i.e., after a collaborator's action is understood, the reason(s) (I or intention) behind that action is questioned in the task context. Both processes have been researched to improve human-robot collaboration. For action recognition, most works require visual information, while intention recognition research has exploited a wider range of data like visual, auditory, language and physiological signal processing. In the following, some of these works are presented.

Deep learning is one of the main approaches in action recognition. Guo et al. [59] reviewed several works based on different deep-learning methods. Despite their high accuracy and wide applicability, these methods require large datasets of labelled visual cues (like skeleton tracking) and not having comprehensive data could lead to training bias. This means the performance of the recognition method highly depends on the image representation. As such, large datasets such as one presented by Ji et al. [68] have been created and used for human action recognition. Poppe [94] has also reviewed several vision-based action recognition research which is mainly based on classification methods like Support Vector Machines (SVM) and Principal Component Analysis (PCA). All these methods require the preprocessing of visual data and reliable detection of the region of interest (ROI) for action detection. Poppe argues that most of these methods also fail when there is a “severe occlusion”. He also points out that a temporal variation of complex actions can be misrecognised for similar actions like jogging and walking.

One could argue that methods used for action recognition like those presented by Parisi [91] using a self-organised deep neural network for action prediction could be equally applied to the prediction of a partner’s decision in a joint human-robot collaboration. While this potentially could be done by mapping said predicted action to a related decision, it is not always possible to directly relate an action to a decision especially when the same action could be the result of different decisions. For instance, the action of picking an object when there are objects with equal priority to be picked up could not accurately represent the decision of which object has been chosen. In addition, such methods normally require a large library of recorded actions (most of the time in the form of a sequence of images) and only can be applied to that specific set of recorded actions.

Similarly, intention recognition research has also utilised deep learning methods and visual information. For example, Cheng et al. [26] use human skeleton tracking data as input to their LSTM network. Along with deep learning methods, other approaches, such as Probabilistic State Machine (PSM) for predicting human intention using skeleton tracking as input [37], gaze tracking [46], using Bayesian models [66, 112], Hidden Markov Models [136] or physiological signals like Electromyography (EMG) [14] are used for intention recognition. These techniques are either probabilistic (as such not desirable as discussed in the next chapter) or require a pre-/post-processing of input data like physiological data and skeleton tracking causing additional complexity for it to be used for decision-making modelling. Nonetheless, intention recognition is an essential step towards understanding of decision process as knowing "the intention of peers" is one of the main inputs to the decision-making process. A clarifying example in a joint human-robot interaction could be when the intention of a human is to pick up a certain object with more than one instance of it involved in the task and a decision need to be made to choose between these instances of the same object. Therefore, in such a scenario like the one in this work, the intention would be picking the object but the decision would be differentiated as a specific instance of it.

2.2 Action Planning

Planning came to be an essential part of robot control in the Sense-Plan-Act model of robot control [83]. For mobile robots, planning involves a search or an optimisation approach to find the best path while avoiding obstacles in the environment and for the robotics arm it entails finding a path to move the end-effector and perform a task [27]. To enable robots to collaborate with human counterparts, other features such as role allocation [85] or task allocation [69], belief management [58], and coordination mechanisms [84] were added to the planning phase. Most of these methods are implemented either through verbal communication or hard-coded rules. Nonetheless, several efforts were made to address the dynamic nature of human-robot interaction. Devin and Alami [34] present a Theory of Mind (ToM) inspired architecture in which a ToM manager is introduced to "estimate and maintain the mental state of each agent" as it receives information from the Situation Assessment module. The main goal of this architecture is to inform the human partner of the progress in the shared plan without being intrusive or annoying if, for example, s/he is distracted during the task. This is a good example of attempts to equip robots with high-level mental models. Buisan et al. [20] also present an architecture for human-aware planning that emulates the partner's decision to better predict his/her actions. In their architecture, a human-aware task planner (HATP) is combined with a simulation module in which the robot's human partner decisions are being predicted and the task planner updates the robot's action based on this prediction.

Overall, to the best of our knowledge, planning research does not address the challenge of dynamic decision-making required for human-robot collaboration. Most work is either based on hard-coded plans or requires verbal communication to resolve conflicts. The recent planning architecture by Buisan et al. [20] has tried to address this shortcoming and is the closest work to the proposed work here. However, the planner does not have the ability to resolve conflicts in parallel actions. It is noteworthy that parallel actions are required to have a natural human-robot collaboration. As explained in the following chapters, in human-human interactions both serial action (turn-taking) and parallel action (working at the same time, were observed during the data collection phases of this work. This is especially important as there could be many steps of the task execution with the same priority and a decision-making module like the one presented in this work is required to enable the robot to work parallel alongside its partner towards task completion instead of waiting for its turn to execute an action.

2.2.1 Planning for Cooperative Object Carrying

As required in human-human interaction, humans and robots might need to engage in a physical interaction in which they carry an object together. This has brought up a challenge of planning and coordination in such tasks and researchers have tried to address it in different ways. Using haptic feedback has been a dominant approach in creating coordinated cooperation between humans and robots [19, 36, 56, 57, 124]. These works require physical contact between the robot and human as they rely on force and torque feedback. A fusion of the visual and feedback has also been used for such tasks [1, 134], Nonetheless, physical contact still is required for such methods to work.

Sheng and his colleagues [106] developed a framework for a humanoid robot to collaborate with a human in lifting a table. They have used an imitation learning approach by using Gaussian Mixture Models (GMM) and applying Gaussian Mixture Regression (GMR) for learning to grasp. The robot then uses two types of controllers either a reactive controller or a proactive controller to adjust the lifting to its human partner and keep the table horizontal. The reactive controller output is based on "the observed state of the object". This controller is trained by reinforcement learning. The proactive controller works like the reactive controller however it acts based on the prediction of the human's next action so that the robot's goal is to be always ready to adjust as the human action is implemented instead of waiting for the action to be done and react to it. To this end, a Kalman filter (KF) is utilised. Such approaches could potentially be extended to be applied to other joint action scenarios like the one presented in this work, however, when training decision-making models, reinforcement learning approaches could require very long training times as the number of alternatives increases.

More recently, Ng et al. [86] proposed a motion planner for mobile robots when carrying an object with a human. The planner is trained based on human-human interaction using a Variational Recurrent

Neural Network (VRNN). The main idea behind their algorithm is similar to the work of Sheng et al. [106]. The planner used a prediction model to adjust the robot's movements to the human's. We also use the same concept, however, we use different modelling methods as explained in the following chapter. In addition, the joint action scenario is different. There are no kinematic constraints imposed by the closed kinematics chain created due to carrying an object together. This means two collaborative partners have a higher degree of freedom and the proposed methods in the literature are not necessarily sufficient for developing predictive models. for instance, the performance of the proposed method is compared to recurrent neural networks and it has outperformed it as it is reported in the following chapters.

2.3 Cognitive Architectures

There are several cognitive architectures developed for individual artificial agents in the literature. Here, we consider three architectures that are deployed in the field of robotics, namely, ACT-R, SOAR and R-CAST.

2.3.1 Adaptive Control of Thought-Rational (ACT-R) Architecture

Developed by John R. Anderson and his colleagues at Carnegie Mellon University, ACT-R (Figure 2.1) aims to provide a framework for understanding how humans perceive, learn, remember, and perform tasks. It is based on a production system, which consists of a set of condition-action rules (productions) [5]. These rules describe how the system should respond to different situations or inputs. ACT-R distinguishes between two primary types of knowledge, namely declarative and procedural. Declarative knowledge contains facts and everything an agent is aware of and can describe to others, while procedural knowledge corresponds to sequences of actions and how to perform tasks and it could be observed in our behaviour and not necessarily conscious. This separation allows ACT-R to model both conscious knowledge and automatic, skill-based behaviours [5]. It operates by setting and pursuing goals. Goals represent desired states or tasks to be accomplished, and the system works to achieve these goals by selecting and executing appropriate productions [4]. As the architecture evolved over the years, it has now a conflict-resolution module for choosing between possible productions. The resolution is achieved by a ranking (or weighting) of the production rules to prioritise their execution [96]. While Trafton et al.

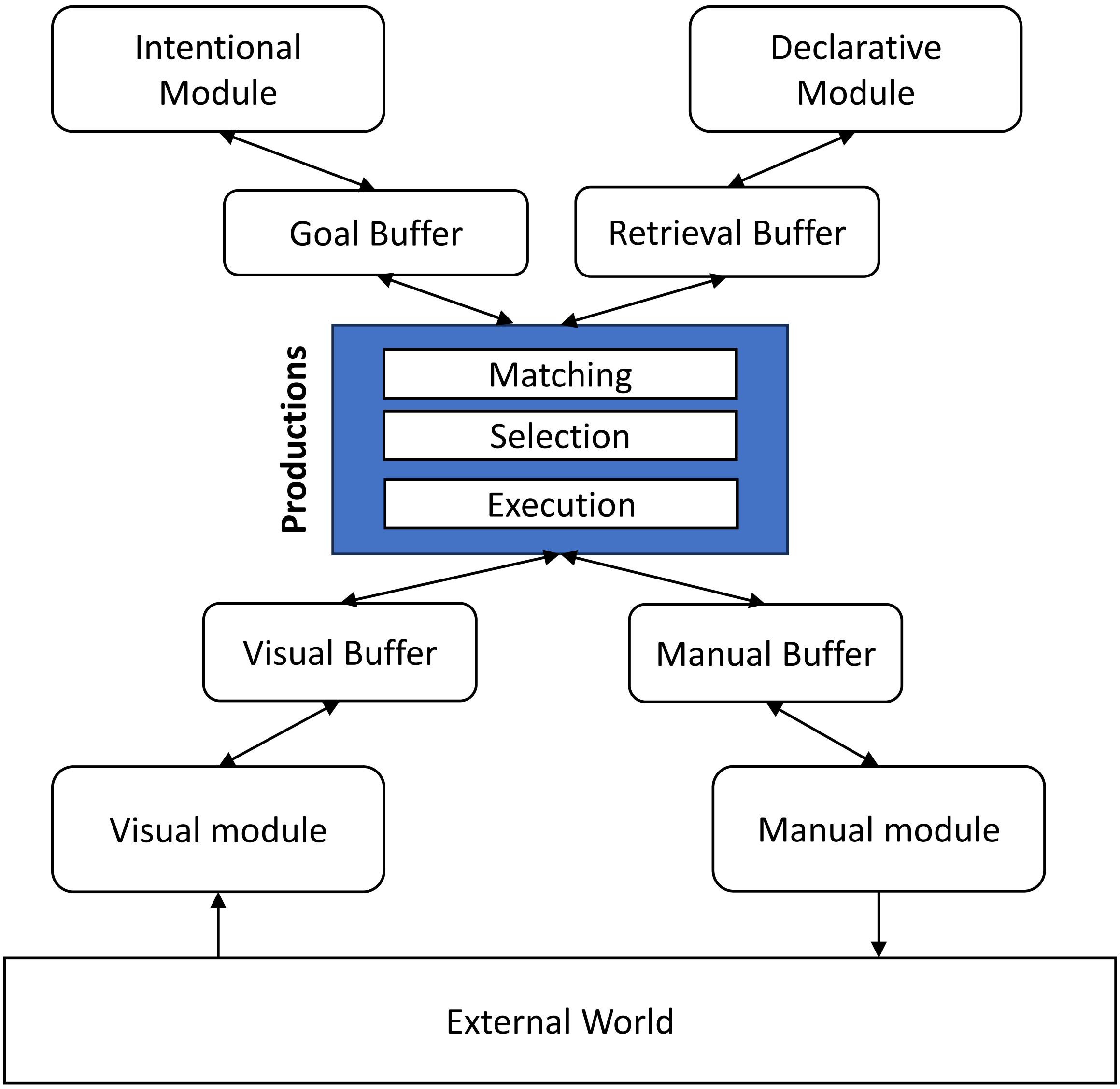

Figure 2.1: Schematic diagram of the ACT-R cognitive architecture. Recreated by the author based on ACT-R 5.0. [4]

[116] have developed an embedded version of the ACT-R, the architecture has never been used for a joint action scenario and lacks a negotiation layer like the one presented in this work.

2.3.2 State, Operator, And Result (SOAR) Architecture

Developed by John Laird and Allen Newell at the University of Michigan, SOAR (Figure 2.2) is designed to model and simulate human cognitive processes and problem-solving abilities. It represents knowledge

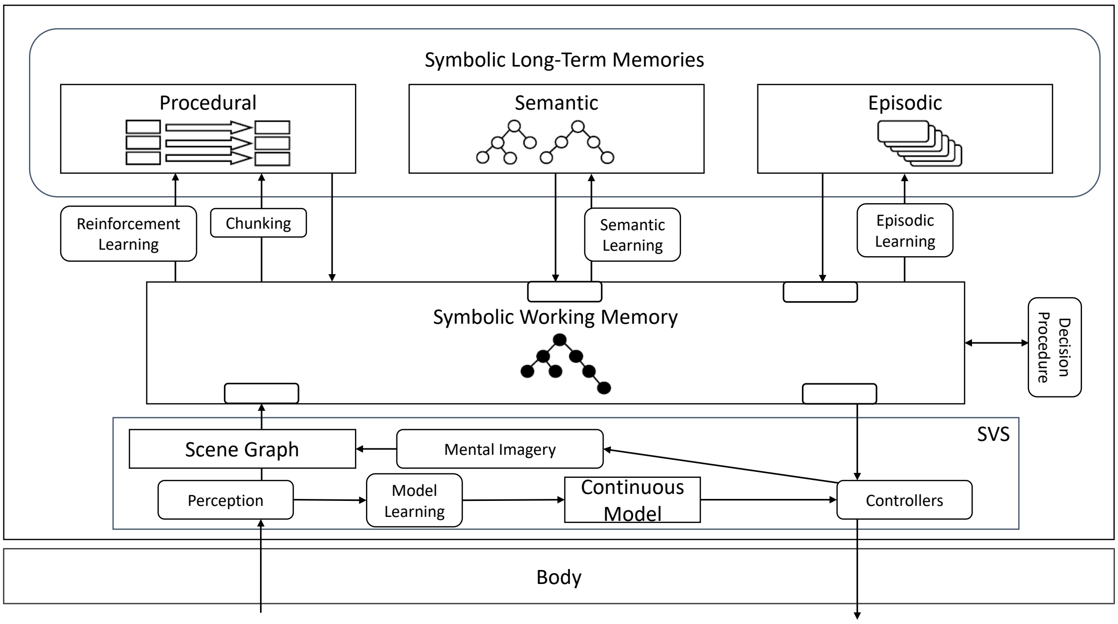

Figure 2.2: Schematic diagram of the SOAR cognitive architecture. Recreated by the author based on [75]

using a production system, similar to ACT-R and consists of a large set of production rules that describe how the system should react to different situations. These rules are written in the form of "if-then" statements [74, 75]. The architecture contains a working memory structure to hold information that is currently being processed. This working memory is highly flexible and can hold a variety of types of information, including symbolic facts, goals, and problem-solving states. It employs a problem-space search mechanism to solve complex problems. It represents problems as states and uses operators to transition between states in search of a solution. To represent and organize knowledge, chunking is used. Chunks in SOAR are units of knowledge that can be created, retrieved, and manipulated during problem-solving and learning. Chunks allow for the efficient handling of complex and structured information [72]. Several learning mechanisms, including reinforcement learning and chunking, are incorporated into the architecture. So, it can learn from experience and adjust its behaviour based on feedback and past performance [72]. In terms of conflict resolution, SOAR prioritises the production rules based on predefined criteria and context [130, 131]. SOAR is used to solve navigation problems [75] that can be extended to mobile robot applications, nonetheless, similar to ACT-R, it has not been used for a joint action scenario. In comparison to ACT-R, having a problem-solving approach through searching the problem space, it could be more capable of addressing joint action conflicts by searching for alternative action. However, the search algorithm could prove computationally extensive as tasks become more complex.

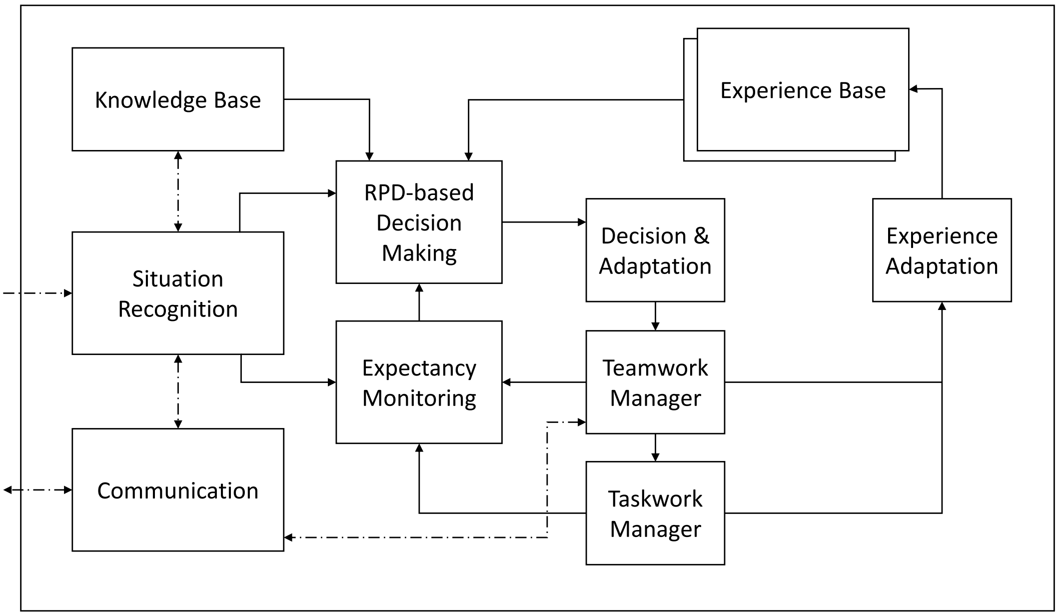

Figure 2.3: Schematic diagram of the R-CAST cognitive architecture. Recreated by the author based on [43]

2.3.3 R-CAST Architecture

Developed by Fan and Yen [43] for human-AI teamwork, it is based on the CAST (Collaborative Agents for Simulating Teamwork) agents [133] equipped with Recognition-primed decision (RPD) models [42]. The R-CAST architecture is depicted in Figure 2.3. CAST agents are designed to simulate multi-agent teamwork based on formal models of information exchange coded in a language called MALLET (Multi-Agent Logic-based Language for Encoding Teamwork). Team processes and structures are hard-coded using MALLET descriptors. The shared team processes are coded using Petri Nets as a "computable model of mental states". Agents' synchronisation is done through Petri Nets by transitions between control nodes and belief nodes [133].

RPD initially was developed as a decision support tool. It operates in two phases, recognition and evaluation. In the recognition phase, the agent uses the information from the situation and experience to

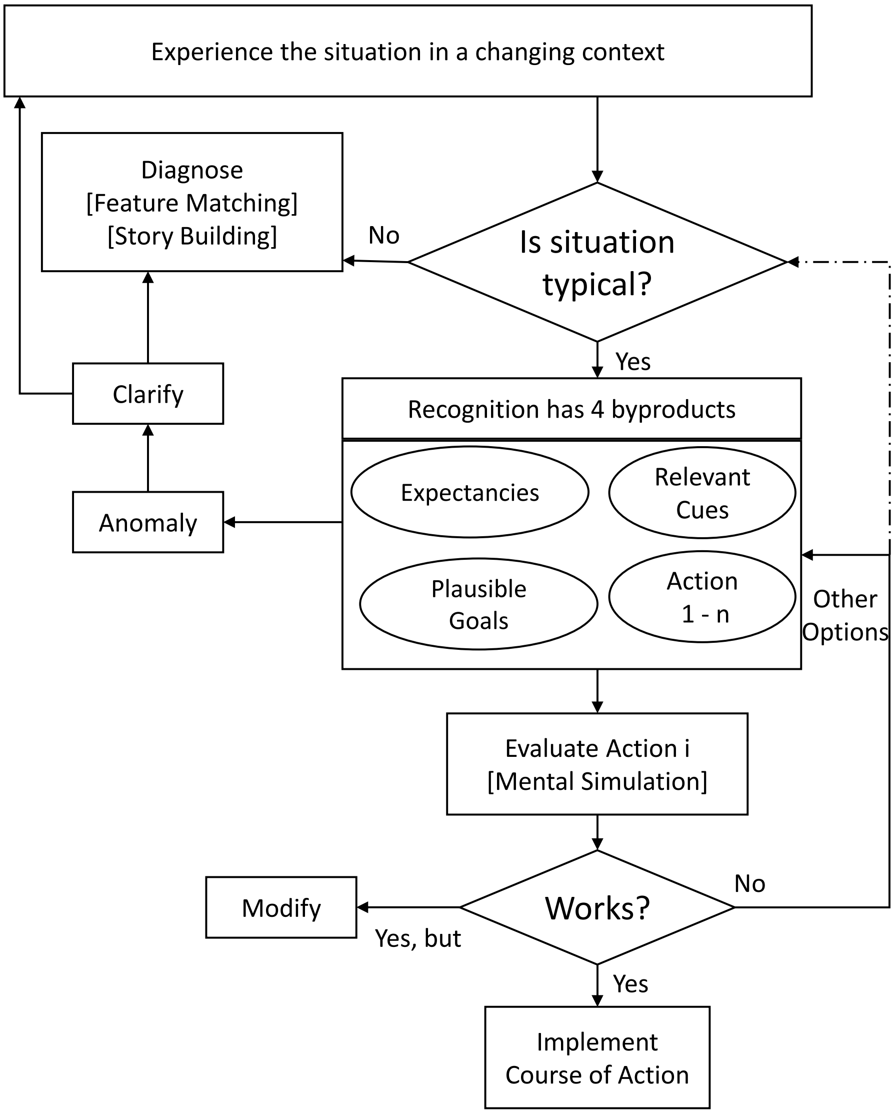

Figure 2.4: Schematic diagram of the Recognition Primed Decision Model. Recreated by the author based on [42]

recognise which course of action has worked in the past in a similar situation. The outcomes of this step are depicted in Figure 2.4 as Relevant Cues (what to pay attention to), Expectancy (what would happen next), Plausible goals (which goal makes sense) and Course of Actions (indicating actions worked before in similar situation). In the Evaluation phase, the decision-making agent would imagine the outcome of a course of action. The evaluation is an iterative process until the agent finds a course of action that works for the current situation. The model also uses two strategies for gaining situation awareness, namely feature matching and story building. These strategies are used by experts first to find whether they ever experienced situations similar to the current one (feature matching) and if not, they will construct an explanation based on linking the observed information (story building) [42]. By merging the RPD model with the CAST agent, the R-CAST architecture was created. In this way R-CAST can first provide a shared mental model in the team and by using the RPD models the decision-making in the team would be more natural as the human decision-making process is entwined with agents' decision-making through tow-way communication [43].

2.3.4 Cognitive Architectures Review Conclusion

Two dominant cognitive architectures namely, ACT-R and SOAR and R-CAST as a multi-agent teaming architecture created based on a decision support tool were presented above. ACT-R and SOAR have provided good insight into different aspects of human cognition. However, the decision-making and conflict resolution in these architectures have limited capacity and would not be the best choice for a dynamic dyadic joint action scenario with parallel actions. The R-CAST architecture is designed for teamwork and could potentially be applied to such a scenario. Nonetheless, the current state of the architecture is based on expert knowledge and explicit communication. Hence, The proposed architecture in the next chapter addresses this shortcoming by equipping the robot with a negotiation layer acting as a means of implicit communication for resolving conflicts.

Chapter 3

Proposed Architecture

Inspired by the role of mirror neurons in joint actions [90] and considering prediction as a crucial part of this process [10, 104], a novel decision-making architecture is designed for HRC scenarios with no explicit communication (either verbal or gesture) between human and the robot. The proposed architecture foresees mental models of the decision-making processes of both the agent and the interaction partner. The architecture is depicted in Figure 3.1. In developing the architecture, it is assumed that the agents involved in the joint action are committed to completing the task and reaching the goal. Agents' intentions are assumed to be known as implementing an action towards completing the task like picking an object, however, their decision of which object to pick is not known. It is also considered that both agents are engaged in the task in a co-active collaboration with equal capabilities.

In the architecture, each agent has its own decision-making system that allows for combining a series of independent individual policies by means of an integration layer. The two decision-making systems of the agents run in parallel when collaborating on a joint action task and their outcome enters a negotiation layer. This layer is introduced to prevent conflicts in action execution by negotiating own independently taken decisions with anticipated partner's decisions. The latter is obtained by internal simulation of the mental model of the partner. So, each agent is assumed to simulate its own and its partner's decision-making process and to integrate the two independent decisions deriving from these processes into one final outcome. The final decision on the next action is reached after both its own and

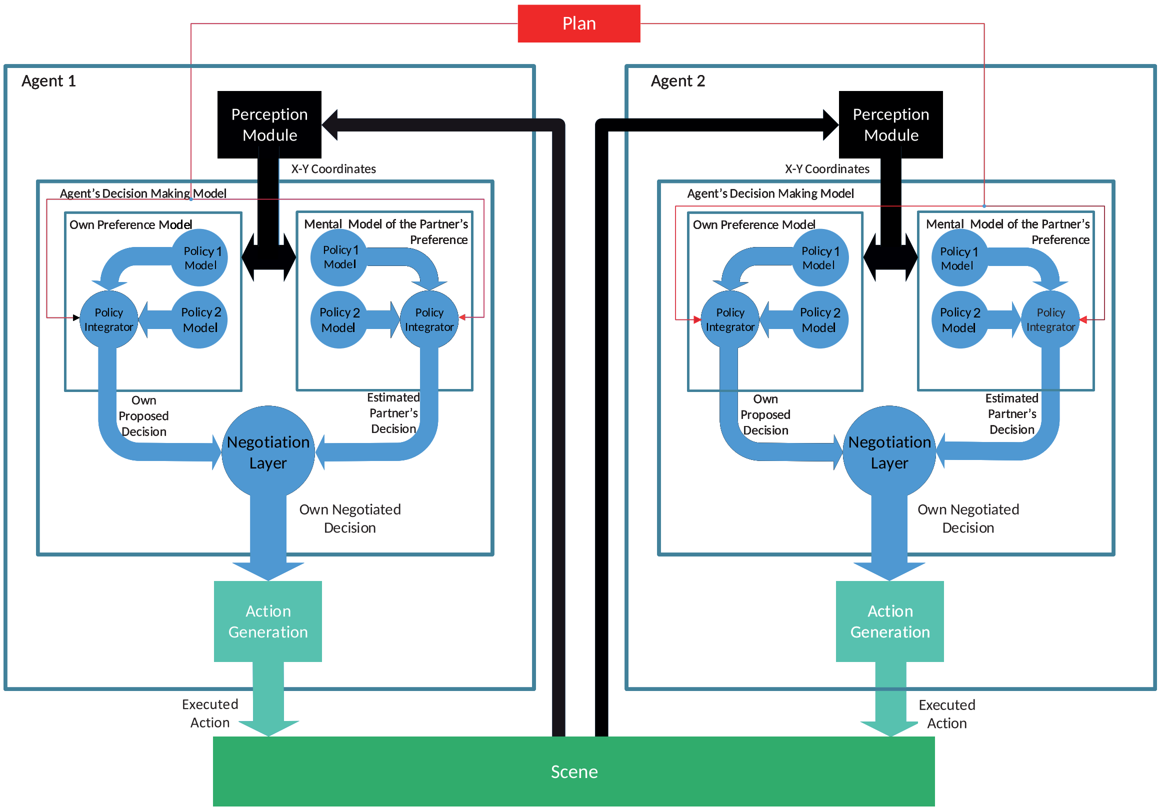

Figure 3.1: Proposed decision-making architecture for joint action: an abstract depiction of the decision-making process of two agents. Each agent's decision-making process takes into account (i) its own preference model (including different policies and policy integrator); (ii) an internal simulation of its partner's decision model (inspired by mirror neurons); and a negotiation layer that combines (i) and (ii).

the predicted partner's decisions are integrated in the negotiation layer. Unlike works such as the one by Devin and Alami [33] that utilises a dialogue system, here the negotiation layer works as an implicit communication, as after the actions of both agents are updated in real-time, these updated actions again trigger a new outcome of the internal simulations of both agents. Thus, if for example, both agents come to the same decision, the one implementing the decision faster will have its action allowed to be executed, while the other will be prohibited to continue until the next foreseen action of the shared plan starts.

In developing the negotiation layer, similar to modelling decision policies, data collected from human-human interaction was used to create a naturalistic conflict resolution. The negotiation layer outcome is updated in real-time. This means as soon as new outcomes from decision models (as a result of any change in the agent's actions) are fed into this layer there will be a different outcome of the conflict resolution process (if there are any conflicts at all). In an extremely rare case in which there is no differentiation measure like approaching the target object faster, the negotiation layer gives priority to the human partner when it is embedded on the robot. However, this was not observed in human-human interaction experiments presented in the next chapter.

It is noteworthy that the architecture is developed as a result of an iterative design in which at first the decision policies were not modelled separately. The first design contained an individual decision model like the one presented in [16], however, this model did not work and the training did not converge after many iterations when trying to have multiple policies like those in the next chapter. Therefore, in the next step, the decision policies were modelled separately and then integrated by the means of decision integration layer leading to the currently presented architecture.

The perception module in the architecture represents any proprioceptive sensors that provide information on body movements as well as sensors that provide information on object movements. In developing the models in this work, a Vicon motion capture system and Microsoft Kinect to track hand and object movements are utilised. However, depending on the complexity of the recognition system, other stereo vision or RGB-D cameras could be used. (The experimental setup and used tracking sensors are described in Chapter 4). While in this work we only consider kinematic motion, there is no constraint imposed for using a dynamic motion and getting haptic feedback data as used in the literature for joint object carrying.

Action generation is considered to be a module that derives required action commands for action execution and is not considered the main topic of this research. It is assumed to be covered by standard motion planning algorithms available in the literature when, e.g., implementing the structure on a robot. It will receive the final decision and generates a series of commands to be sent to the low-level control system of the actuators that then execute the individual actions.

The plan module is to implement a shared plan for joint action and differentiates between parts of the plan that can be executed in series or also in parallel. It also activates the related policy models or layers required for task implementation.

The presented architecture is designed for dyadic joint action. The models presented in the following chapter are also developed based on dyadic human-human interaction and the models and architecture are tested later on in a dyadic human-robot joint action. The Negotiation layer in the architecture works in real-time and can update and adapt to any changes in the interactions. Nevertheless, while modelling decision policies could be applied to any number of interactive agents, as the decision needs to be updated in real-time, employing the negotiation layer of the architecture for any higher dimensional interaction of triadic or more requires rigorous testing beyond the scope of this thesis.

3.1 Selection of Mathematical Modelling Framework

Considering the dynamic nature of human decision-making processes, the desired mathematical framework for modelling the decision-making module has to be able to implement this dynamic and predictive nature of the processes. Dynamic methods can model the high variability of a system over time. For decision-making, a dynamic decision model can address the change of decision due to different factors like environmental events. Although there are many dynamic probabilistic modelling approaches available in the literature, finding accurate probability information on human decision-making would require a large database. This is due to human behaviour being affected by several internal and external factors [95], a decision-making process could have different outcomes in different situations, hence, a probabilistic model requires all these varied decisions (a large dataset) to be trained. This makes the modelling based on such approaches difficult, if not impossible as it would require recordings of a large series of real human-human collaboration experiments. Furthermore, while many probabilistic approaches have been used successfully when the required computational resources are available, many "naturalistic decision-making researchers argue that when people make a decision in their day-to-day life actions they do not know the probabilities of all alternatives and sometimes they might not even know all possible choices [108]. Thus, in this work deterministic methods are preferred (not considered superior to probabilistic approaches) as then modelling requires a relatively smaller dataset. As argued by Kahneman and Tversky [70] people do not necessarily make rational decisions. So, to have a well-generalizing model, methods based on rationality assumptions (the tendency to maximise utility when making a decision) are not suitable for this work. At the same time, the system should be able to cope with uncertainty and multiple alternatives (considering most decision-making models focus on two-alternative forced choice) while avoiding assumptions hence avoiding normative (in which only one best decision exists from the choice alternatives) approaches. So, the features of the required modelling are categorised as:

- **Desirable**: Predictive, Multi Alternative, Dynamic, Coping with uncertainty
- **Undesirable**: Probabilistic, Normative, Static, based on Rationality assumption

To finally choose a proper mathematical modelling framework, some of the well-known techniques for implementing decision-making were reviewed as presented in the following paragraphs.

3.1.1 Decision Trees

Decision trees are one of the most popular decision support tools, a tree-like graph that starts from the decision that needs to be made and branches to the chance nodes and further sub-decisions and consequences of the decision in different situations. The value of uncertain outcomes O is calculated by multiplying the probability by the gained value of O. Each Decision node could have several chance nodes. The value of each chance node is calculated by finding the maximum net value of outcomes branched out from that node. the net value of each outcome is calculated by reducing the cost of the outcome from its value [60, 89]. For instance, the value of the chance node in the decision tree depicted in Figure 3.2 is calculated as presented in the following equations.

$$V = P * V_O \tag{3.1}$$

where V is the value of end node with outcome O with probability of outcome P. Then, the value of chance node C_1 is calculated as

$$V_{C1} = MAX\{V_1 - C_{O1}, V_2 - C_{O2}\} \tag{3.2}$$

where V_{C1} is the value of chance node C_1, V_1 and V_2 are value of end nodes with outcomes O_1 and O_2, finally, C_{O1} and C_{O2} are costs associated with outcomes O_1 and O_2.

Decision Trees (DT) have been used in many different fields, for example: i) corporate decision making; ii) Artificial Intelligence (AI) and machine learning for applications like decision support, regression, data mining; iii) path planning for mobile robots [63, 111]. There has been an effort to make

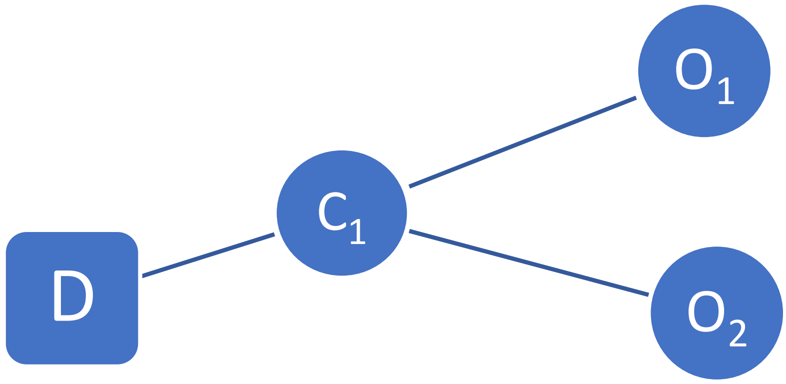

Figure 3.2: Part of a Decision Tree with decision node D, chance node C_1 and two possible outcomes O_1 and O_2.

DTs as dynamic as possible [63]. However, they have been generally found to be not applicable when decisions have to be made for a continuously changing and dynamic environment. This is due to the fact that the general structure of the tree and the main consequences, including their probabilities of occurring, have to be known at the outset. This is not always possible for applications like HRI when human behaviour needs to be considered when establishing the DT structure.

3.1.2 Expected Utility and Prospect Theory

In classical economics, Expected Utility (EU) Theory is used in a descriptive way trying to explain *why* people make a specific decision. In philosophy, on the other hand, it is used as a normative theorem explaining how people *should* make decisions. The essence of the theory is that people are considered rational so they will make decisions to maximise the utility of the outcome of their action [62]. The action of the decision maker will be state-dependent and, since the states are uncertain, the expected value is calculated as a probabilistic weighted sum of the utility of outcomes of action in different states.

So, expected utility of action A is calculated as

$$EU(A) = \sum_{O=1}^{N} P_A(O)U(O) \tag{3.3}$$

where $P_A(O)$ is the probability of having outcome O when choosing action A and $U(O)$ is the utility of outcome O [62]

This theory has been popular in different disciplines to explain human decision-making. However, as will be explained in the following, EU theory has difficulties predicting human behaviour. In terms of its use in robotics, there is research on action planning using utility maximisation like [98] with reported improved performance of planning. Like DT this approach relies on knowing the probabilities of consequences and needs information on the task at the outset.

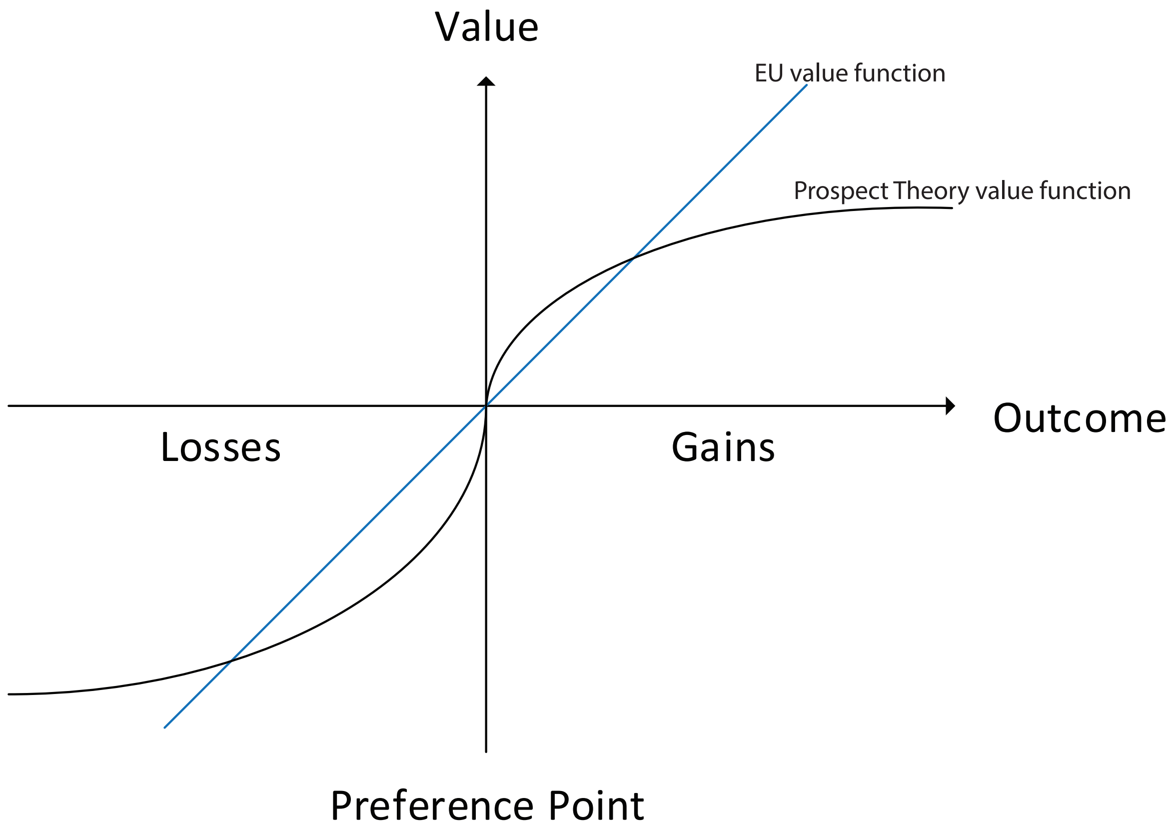

Figure 3.3: Value functions for EU and Prospect Theory.

As Kahneman and Tversky [70] well pointed out, EU theory, as a descriptive or predictive theorem, is likely to fail when it comes to real-life decision making. Instead, they suggested Prospect Theory which tries to explain why people are not always rational and do not always make optimal decisions. The main idea of this theory is that people are neither always risk-averse nor always risk-seeking. They mostly seek risk when there is a high loss and mostly avoid risk when there is a high gain. This makes the value function, describing the value of an outcome, nonlinear in contrast to the linear one in EU theory. Having a steeper value function for losses means they have a higher effect than gains. Hence, in Prospect Theory the final utility gets lower as people give less value to higher gains by avoiding and not taking risks in such a situation (Figure 3.3). Conversely, when there is a high loss people tend to take higher risks and they give a higher value to the outcome compared to what seems to be the rational value. Like Utility Theory, Prospect Theory has been used in many disciplines to explain human behaviour and decision-making processes. Particularly in robotics, for example, it has been used to model human behaviour for assistive robots [120]. Prospect Theory is also relying on knowing probabilities of events and consequences which limits its use in highly dynamic environments. In addition, Expected Utility and Prospect Theory have been mainly used for two alternative tasks but increasing the number of alternatives may render the problem highly complicated.

3.1.3 Markov Decision Processes (MDP)

MDPs are a mathematical discrete stochastic model of decision-making. An MDP includes several states, in each of which the decision maker can choose from a pool of available actions. The probability of moving from one state to another is a function of the current state, so, the next state depends on the current one and the chosen action by the decision maker. The decision maker will receive a reward each time the process moves from one state to another [11]. The whole process relies on having complete knowledge of finite states and actions. MDPs have been used in several applications like economics, automated control, manufacturing, and robotics.

A more generalised variant of MDPs are Partially Observable Markov Decision Processes (POMDP) in which the process does not have complete information on the current state (the current state is uncertain) and not all the states are completely known or "observable". POMDPs use probability distributions to represent how the environment evolves over time. These transitions are typically modelled as Markovian, meaning that the future state depends only on the current state and action taken [23]. Hence, similar to MDPs the transition to the next state is a function of the current state and current action. The main goal of both MDPs and POMDPs is to maximise the cumulative reward by optimising the policies for choosing actions. So, the reward is maximised as

$$Max\{\sum_{t=1}^{\infty} \delta^t R_{a_t}(S_t, S_{t+1})\} \Rightarrow \exists a_t = \pi(S_t) \tag{3.4}$$

where R_a is the reward received due to the transition to the next state and is a function of the current state and current action. δ is a discount factor indicating the importance of future and present rewards and its value is between 0 and 1, and π is the policy based on which the action a is chosen.

Among all the probabilistic approaches, POMDPs have been used most in robotics as they can be applied in uncertain and dynamic environments. Examples can be found in control, planning, and navigation [45, 92, 93, 109]. POMDPs have been applied to a vast range of fields like machine vision, business, corporate policy, and marketing. However, they can only deal with problems with certain characteristics such as having a finite state set and following the Markov Property (meaning that future states only depend on the current state and not past states). Also, it can be highly computationally expensive to assess all the rewards, transition probabilities, and observation probabilities [22] and thus, solutions are often approximated. In addition, the data collected in the human-human interaction phase of this work was used by [125] to evaluate the feasibility of using POMDP as the modelling approach. However, the POMDP models completely failed. This is likely due to the noise in the data as the recorded data consists of tracking coordinate frames of participants' motion and there are many short temporary losses of tracking.

3.1.4 Decision Field Theory

Decision Field Theory was introduced by Busemeyer and Townsend [21], as a dynamical stochastic mathematical model of decision-making, initially focusing on problems of approach-avoidance behaviour [114]. In contrast to normative theories, it tries to explain people's behaviour and decisions without a rationality assumption. The main feature of this theory is that it dynamically models the evolution of the decision during deliberation time rather than considering fixed states of preference. The theory is based on two main psychological principles namely, approach-avoidance in motivation theories and information-processing theories of choice response time [21].

Busemeyer and Townsend developed DFT in an incremental way starting from basic deterministic *Subjective Expected Utility (SEU)* theory and building upon this by adding processes in 7 stages moving from deterministic SEU to *Random SEU*, *Sequential SEU*, *Random Walk SEU*, *Linear System SEU*, *Approach-Avoidance Theory* and finally presenting Decision Field Theory (DFT). These stages are explained briefly here.

3.1.5 Subjective Expected Utility (SEU)

Considering two alternative tasks and using the probabilistic weighted sum to calculate the utility of each action as in the above-mentioned Expected Utility Theory. a new parameter d, mean difference, is introduced. The mean difference is computed by subtraction of the calculated utility of actions which is $d = EU(A_1) - EU(A_2)$. For two alternative tasks if d > 0 it means the preference direction is towards A_1 and vice versa.

3.1.6 Random SEU

The next step is the development of Random SEU in which the decision maker has the freedom of switching between choices across trials. So, according to Random SEU, there is no fixed probability of outcome and rather the *attention weight* is changing from one trial to another resulting in the SEU of each action being a random variable which is called *valence of action*, V. It is calculated in the same way as EU but rather than having constant P_A it will be a continuous random variable denoted W to model fluctuation of attention. So,

$$V(A) = \sum_{O=1}^{N} W_A(O)U(O) \tag{3.5}$$

The difference of valence of actions, P, determines the choice similar to mean difference d. For two alternative tasks $P = V(A_1) - V(A_2)$. And if P 0 then A_1 is the chosen action and vice versa.

The difference between SEU and random SEU is presented by residual difference, $\varepsilon = P - d$ which shows the change of preference from trial to trial. So, $P = V(A_1) - V(A_2) = d + \varepsilon$ which means to choose action 1 over 2, Pr[P>0] $= Pr[\varepsilon \;\; -d]$ which is a representation of the mathematical model of random SEU. It is assumed that ε is normally distributed by zero mean and $Var(\varepsilon) = Var[V(A_1) - V(A_2)] = \sigma^2$

which is called the *variance of the valence difference* and is used to show the strength of preference. By considering F as a normal cumulative distribution function we have,

$$Pr(A_1,A_2) = Pr[\varepsilon \quad -d] = F[(\frac{d}{\sigma})] \tag{3.6}$$

meaning the choice probability is an increasing function of $(\frac{d}{\sigma})$ which is called the *discriminability* index. From basic statistical theory

$$\sigma^2 = Var[V(A_1) - V(A_2)] = \sigma^2_{A_1} + \sigma^2_{A_2} - 2\sigma_{A_1A_2} \tag{3.7}$$

Where $\sigma^2_{A_1}$ is the variance of the valence for action 1 and $\sigma^2_{A_2}$ is the variance of the valence for action 2 and $\sigma_{A_1A_2}$ is covariance of these two valences which is negatively related to the variance of valence σ^2. This means by increasing the similarity of payoffs for actions the discriminability index will increase[9].

The shortcoming of random SEU is it cannot provide an explanation of the systematic relation of decision time and choice probability. This is addressed in the next stage by going development of the Sequential SEU theory.

3.1.7 Sequential SEU

To allow integration of decision time in the Sequential SEU there will be a sequence of one or more samples during deliberation time of each trial meaning that the decision maker preference may even change in each trial in addition to one trial to another. So, the preference state will be accumulated over all the samples. This means if $P(1) = [V_{A_1}(1) - V_{A_2}(1)]$ then $P(2) = P(1) + [V_{A_1}(2) - V_{A_2}(2)]$ and so on. So, for $n \quad 2$:

$$P(n)) = P(n-1) + [V_{A_1}(n) - V_{A_2}(n)] = \sum_{k=1}^{n} V_{A_1}(k) - V_{A_2}(k) \tag{3.8}$$

$P(n-1)$ is the previous preference state of having $n-1$ samples and then $V_{A_1}(n) - V_{A_2}(n)$ is the new valence difference. This process will continue until it reaches a threshold θ which is called inhibition threshold. In the case of having a positive preference state, the preference is in favour of action 1 and the action will be chosen as soon as the accumulative preference reaches the inhibitory threshold θ and for negative preference, action 2 will be the outcome of deliberation. This means that the decision time is a function of the number of samples required to reach the threshold. Similar to random SEU the residual here represents the change in preference state due to attention fluctuation in deliberation time, so,

$$P(n)) = P(n-1) + [d + \varepsilon(n)] \tag{3.9}$$

Then the probability of choosing action 1 over 2 will be calculated by [30]

$$Pr(A_1,A_2) = F[2(\frac{d}{\sigma})(\frac{\theta}{\sigma})] \tag{3.10}$$

Where F is the standard logistic cumulative distribution function:

$$F(x) = \frac{1}{1+e^{-x}} \tag{3.11}$$

Compared to random SEU, both theories describe the probability of choosing actions as a function of the discriminability index ($\frac{d}{\sigma}$), while in sequential SEU this probability is also a function of threshold θ. In extreme decision cases, the discriminability might be constant or very low so that an increasing threshold causes a high probability of choosing the action with a higher SEU. However, the cost of increasing the threshold θ is having a higher number of samples to reach the threshold which means a higher decision time. The mean number of samples for getting to the threshold level is calculated by

$$E(N) = (\frac{d}{\sigma})[2Pr(A_1,A_2) - 1] \tag{3.12}$$

The downside of sequential SEU is that it cannot explain having choice probabilities of less than 0.5 and for any positive mean difference value the choice probability is always greater than 0.5. So, in the next step, this is addressed in Random Walk SEU theory.

3.1.8 Random Walk SEU

In Random Walk SEU unlike Sequential SEU, the initial preference is not considered 0. It means the decision maker may have a preference based on past experiences and recalling memories which causes the mean difference to lean towards one choice. This generalises sequential SEU theory by adding some anchor point, z, for initial preference $P(0)$. So by having $P(0) = z$,

$$P(n)) = P(n-1) + [d + \varepsilon(n)] = z + \sum_{k=1}^{n} V_{A_1}(k) - V_{A_2}(k) \tag{3.13}$$

The mean difference, d, and variance of the valence, σ^2, are calculated similarly to sequential SEU theory. A good example of choice experience is when the decision maker is choosing between two options, one is known and another one is new. There is an initial preference towards the well-known choice and under time pressure it is more likely that it will be chosen. However, with longer deliberation time it is likely that a positive mean difference causes the decision maker to switch preference towards the new choice if it has some advantages, like higher quality. This models the effect of time pressure on decision-makers.

The inadequacy of random walk SEU theory is considering the effect of primacy and recency in decision-making as the final preference is only sum of initial preference and all the valence differences and it does not depend on the order of the events in sequence. So, the Linear System SEU theory is introduced to tackle this issue.

3.1.9 Linear System SEU

Linear System SEU theory incorporates the effect of the order of events (if that happened late or early) on final preference. All the parameters at this step are the same as random walk SEU theory except new parameter, s, which is called the growth-decay rate. The final preference is then calculated as

$$P(n)) = (1-s)P(n-1) + [d + \varepsilon(n)] = z(1-s)^{n-k} + \sum_{k=1}^{n} V_{A_1}(k) - V_{A_2}(k) \tag{3.14}$$

This makes the random walk SEU theory a specific case of linear system SEU theory by setting s as zero. With this formulation, for growth-decay rate values between 0 and 1, there will be a recency effect on final preference meaning that the later or more recent events have a higher influence and if s is less than zero there will be a primacy effect meaning that earlier events have more influence on the final preference.

The shortcoming of the linear system SEU theory, however, is being neutral on the nature of the decision. The final preference depends only on the sum of valence differences and the well-established findings that the mean deliberation time is longer for avoidance-avoidance problems than approach-approach ones when having constant mean differences. To address this issue, the Approach-Avoidance Theory is introduced in the next step [114].

3.1.10 Approach-Avoidance Theory

In approach-avoidance theory, the aforementioned problem is addressed by adding another dynamic parameter to the equation for calculating final preference. This new parameter is called goal gradient and it implements the effect of the attractiveness of rewards or aversiveness of punishments as a decreasing function of distance to implementing an action [114].

According to this theory, the closer the decision-maker gets to choosing an action, the more s/he will pay attention to the action's consequences. So, consequences of action with little chance of being chosen will more likely be ignored, while when getting closer to the inhibitory threshold for an action its consequences will be more noticeable to the decision maker.

Considering a and b as the goal gradients for rewards and punishments,

$$d(n) = [EU_{A_1}(n) - EU_{A_2}(n)] = -cP(n) + \delta \tag{3.15}$$

where

$$c = b[EU_P(A_1) + EU_P(A_2)] - a[EU_R(A_1) + EU_R(A_2)] \tag{3.16}$$

$$\delta = [EU_R(A_1) + EU_R(A_2)](1 - a\theta) + [EU_P(A_1) + EU_P(A_2)](1 - b\theta) \tag{3.17}$$

and EU_P is the average punishment (loss) and EU_R is the average reward (gain) and by imposing this new mean valence difference to linear system SEU equation:

$$P(n)) = (1-s)P(n-1) + [d(n) + \varepsilon(n)] = [1-(s+c)]P(n-1) + [\delta + \varepsilon(n)] \tag{3.18}$$

δ, mean valence input, in this equation, is similar to d, mean difference, in linear system SEU theory. A positive δ means that the preference is moving towards the positive direction on the average and vice versa. The residual $\varepsilon(n)$ has a zero mean and its variance, σ^2, is calculated similarly to linear system SEU theory. The actual new variable in the approach-avoidance theory is the goal gradient parameter, c. By setting this parameter to zero the above equation will be equal to the linear system SEU theory equation. c 0 defines an avoidance-avoidance conflict and for approach-approach conflicts, c is negative which leads to faster reaching the threshold and shorter decision time. The next step then will be to move to a continuous/real-time approach rather than the above discrete-time methods to form Decision Field Theory.

3.1.11 Decision Field Theory

Building upon previous theories, Busemeyer and Townsend [21] developed Decision Field Theory (DFT). DFT is considered both dynamic and continuous in time in contrast to previous theories which are discrete in time. The real-time DFT is defined by introducing a time variable h which is the time needed to process each sample of valence difference. This is equal to the time needed to process a pair of predicted consequences before switching attention to another pair. Then, the deliberation time, t, will be the total time of processing all the samples consequences or $t = nh$ where n is the total number of samples. So,

$$P(n)) = [1-(s+c)h]P(t-h) + [h\delta + \varepsilon(t)] \tag{3.19}$$

where $\varepsilon(t)$ is the residual input with zero mean, and $h\sigma^2$ variance and σ^2 is calculated same as linear system SEU theory. By having h approaching zero the preference state will be developed in a roughly real-time way. In summary, 7 parameters govern DFT, δ, mean valence input, shows the preference direction, σ^2, the variance of the valence, shows the strength of preference, θ, inhibitory threshold, explains speed-accuracy trade-off, z, the initial anchor point, describes preference reversal as a function of time, s, the growth-decay rate, is used to remove serial positioning effect, c, the goal gradient variable is to explain nature of approach-avoidance conflict during the deliberation time and zero approaching h, as a time unit, is to estimate a real-time process. Although the initially presented formulation of DFT is for a two-alternative task, a multivariate DFT has been also presented in a connectionist interpretation way [97].

DFT has been applied to different cognitive processes like visual sensory detection [107] and conceptual classifications [87]. However, since DFT is in its nature a Markov process [21], Markov assumptions are assumed to hold. In contrast, human decisions may depend on past experiences. The main advantage of DFT over other decision-making models, however, is that DFT tries to explain the *process* of decision-making rather than merely the end result, as it models the evolution of the decision during the deliberation time.

3.1.12 Dynamic Field Theory

Dynamic Field Theory was introduced by Schöner [102] based on the mathematical formulation of dynamic neural fields by Amari [3], as a framework for modelling cognitive processes like detection, selection or working memory. It combines the dynamics of attractors and repellers to form a dynamic behaviour, formulated as follows:

$$\tau\dot{u}(x,t) = -u(x,t) + h + S(x,t) + \int w(x-x')\sigma[u(x',t)]dx' \tag{3.20}$$

$$\sigma(x) = \frac{1}{1+e^{-\beta x}} \tag{3.21}$$

where τ is the time scale, u the activation function over the feature space x at time t, h 0 a constant resting level, S an external input or stimulus to the field, and the integral part is to drive lateral interaction in the population with $w(x-x')$ as interaction kernel and $\sigma[u(x',t)]$ a sigmoidal nonlinear threshold function with a scaling parameter β. Depending on the type of interaction kernel, the nature of the interaction can change from global inhibition to local inhibition or global excitation. This property is being used to model different cognitive processes. A global inhibition, for example, is used for the selection process to achieve a stable choice with minimised effect of environmental noise on the process so that unless the target is shifting to another alternative, the choice won't change due to small environmental perturbance, or a local inhibition is required for a detection process in which the neural field needs to be inhibited in the immediate vicinity of the point of interest so that it stands out of the neighbouring points. These are achieved through the interaction kernel as depicted in Figure 3.4. As can be seen from the equation, the interaction is computed through convolving the interaction kernel with a sigmoidal threshold of the activation function.

The lateral interaction kernel is the key player in changing the behaviour of the dynamic model. One common formulation of the kernel is the following exponential equation:

$$w(x) = c_e x^{\frac{-x^2}{2\sigma_e^2}} - c_i x^{\frac{-x^2}{2\sigma_i^2}} \tag{3.22}$$

where subscript e stands for the excitatory and i for the inhibitory part of the kernel. By changing values of c_e and c_i the excitatory or inhibitory effects of the kernel can be varied. Similar to a normal bell curve σ_e and σ_i are to adjust the bell shape. In Figure 3.4 three cases of interaction kernels are depicted, the green curve is for modelling working memory, the blue one is to model a detection mechanism and the red one is to model a selection process in the neural field.

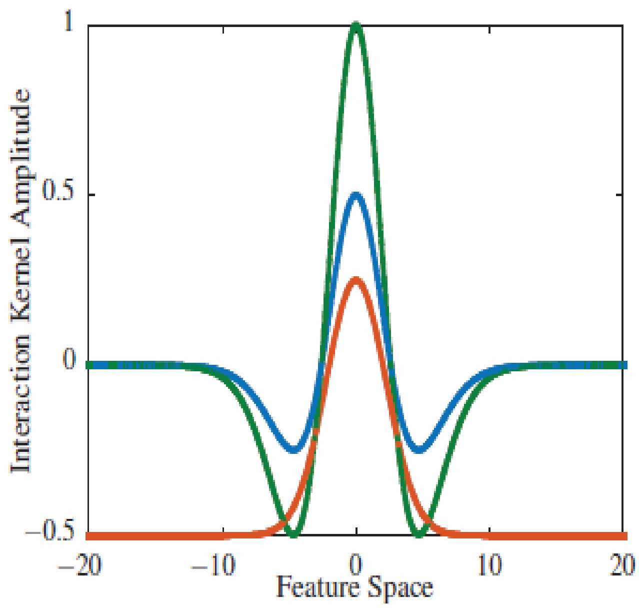

Figure 3.4: Examples of interaction kernels: the green curve is for modelling working memory, the blue one is to model a detection mechanism and the red one is to model a selection process.

Neural Field Theory has been used in the field of cognitive science to model sensorimotor decisions [126], visual cognition [49], modelling object localisation in the visual cortex [67], modelling visual perception [38] and action understanding [40]. In terms of robotics applications, Dynamic Field Theory has been applied to areas like navigation [103], aspects of human-robot interaction and collaborations like action understanding through imitation [41], object recognition [44], verbal and non-verbal communication [17], decision making and joint action for human-robot collaboration [16, 39].

Table 3.1: Comparison of Decision-Making Modelling methods. (red: undesirable, green: desirable, ✓: feature available, X: feature unavailable)

	Decision Tree	EU	Prospect Theory	POMDP	DFT	DNF
Probabilistic	✓	✓	✓	✓	✓	X
Normative	✓	✓	X	X	X	X
Predictive	X	X	✓	✓	✓	✓
Multi Alternative	X	X	X	✓	✓	✓
Dynamic	X	X	X	✓	✓	✓
Coping with uncertainty	X	X	X	✓	✓	✓
Rationality Assumption	✓	✓	X	X	X	X

3.1.13 Summary of Comparison of Mathematical Frameworks

Characteristics of mathematical modelling frameworks introduced so far are summarized in Table 3.1. In this table, when a method has a feature, it is shown by a ✓, and if it doesn't with an X; colours red and green mean undesirable and desirable, respectively. Apart from Dynamic Field Theory, the reviewed approaches are mainly probabilistic methods requiring information on the probability of actions and consequences. While Expected Utility theory-based methods require static and completely known problems in terms of alternatives and consequences, Markov Decision Processes, particularly POMDPs, Decision Field Theory, and Dynamic Field Theory are applicable to uncertain and dynamic problems. Most of these approaches have been initially developed for two-alternative forced choice tasks, however, some like POMDP, DFT, and Dynamic Field Theory (DNF) can be extended to multivariate alternatives. As can be seen from Table 3.1, Dynamic Neural Field (DNF) has all the desired requirements for developing a dynamic decision-making model, even when only small datasets are available. In addition, most of the methods are based on Markov property and are known as "memoryless". However, using DNF a memory trace can be modelled for learning as well the dynamic evolution of the current output of the dynamic field always is a continuous change. This means having a different sequence of actions could lead to a different field activation and hence outcome in each trial.

CHAPTER 4

DEVELOPING MODELS

To model decision policies and develop the initially-introduced decision-making modules, a series of experiments are designed to have either human individuals or dyads to work on an instructed task so that their behaviour could be observed and data required for modelling their decision processes could be recorded. Two main policies are chosen, namely *Distance policy* and *Colour policy*. The former represents an unconscious cognitive process in which people try to minimise their energy consumption and the latter is for the conscious cognitive process of learning plans order. These policies are then integrated using the policy integration layer. The model development steps are presented in the following. Ethics approval for these experiments was obtained from the ethics committee of the University of the West of England (reference number: UREC16-17.03.10).

4.1 Experimental Setup

Participants were asked to work together in a table-top pick-and-place task. Participants were monitored and a set of data consisting of tracked 3D hand and chest position were recorded using a Vicon motion capture system. The motion capture worked with a sample rate of 100Hz, and initially all its data was recorded to create the training datasets. However, the number of samples was too high and this made the optimisation process (explained in the following sections) very slow, therefore, data was down-sampled by a factor of 15. No other post or pre-processing was performed on the data. In addition, at the start of the experiment, participants had their tracked hands resting on the table in front of them. To have a clear baseline for evaluation, participants were instructed to perform the task in a particular way, following a specific pre-defined policy as introduced further below. The blocks were equipped with Augmented Reality (AR) markers and their motion was captured using a Kinect sensor. The experimental setup is depicted in Figure 4.1.

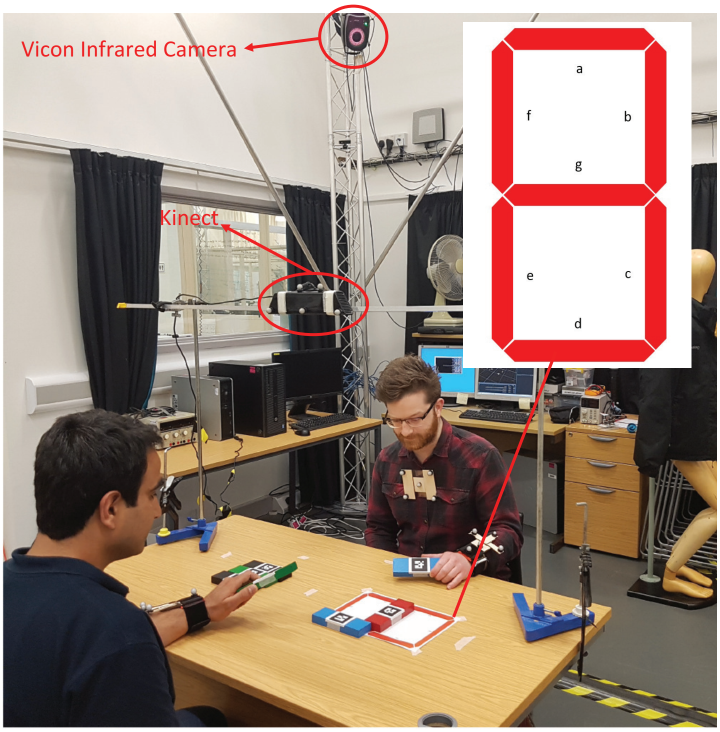

Figure 4.1: Experimental setup. Two subjects collaborating in the dyadic condition. The 7-segment shown on the top right is placed on the table horizontally so that segment "d" comes to lie at the center of the table. The subjects have markers attached to their hands and chest that are captured by the Vicon tracking system. The AR tags on the blocks are tracked by the Kinect camera.

4.2 Task

The chosen task was designed considering certain requirements. The task was supposed to capture an aspect of day-to-day life and be able to be completed by either individuals or pairs. It should be of an abstract level and be extendable or generalisable later on to more complex tasks. Also, as the focus of the work is on the process of decision-making, the task should be as simple as possible to not require any other cognitive processes like problem-solving adding cognitive load, which might affect participants' decision-making.

Taking these considerations into account, the task was chosen as follows: Participant(s) were asked to use provided coloured blocks to form some alpha/numeric characters on a 7-segment template, according to a provided instruction. In terms of experiments with an individual participant, each person was asked to pick blocks one by one and place them on the marked 7-segment shape. In the case of joint action, each participant was asked to pick blocks one by one, being told to "work together", but leaving them the freedom to either take turns or pick blocks at the same time.

4.3 Instructions and Procedure

Before starting the experiment, participant(s) were made familiar with the setup and handed out an information sheet that described their task. They were asked to sign a consent form and were informed that they can withdraw their participation at any point. Then, the participants' right/left wrist and chest were marked with motion capture markers. In each trial participant(s) were asked to perform and complete characters "H", "3", "E", "9", "6", "2", "5", and "8" on the 7-segment shape using provided coloured blocks. For instance, a number 9 could be formed by covering segments"a", "b", "c", "d", "f" and "g". These characters were chosen to counter-balance any effect of the blocks' final position on the outcome (e.g. H and 8 are symmetric while others are mirrored with respect to different axes). To counter-balance any effect of blocks' initial positions, blocks were randomly placed in the middle of the table by the experimenter after each character was formed, with initial orientation changing between vertical or horizontal placement (Figure 4.2). Blocks' in-the-line position was also randomised. Participants were asked to sit on either side of a table with minimum possible movement to perform the calibration of all tracking systems assisted by the experiment conductor. Participants were provided with a set of blocks marked with AR markers (Augmented Reality markers) and were instructed to grab each block in a way not to cover the markers. Participants were also asked to pick and place blocks one by one and to follow the predefined policy. In doing so, participants were instructed to work together without any verbal communication, gesture or facial expression. Participants' movements were recorded when performing the task. In terms of the collaborative phase, participants were asked to always place

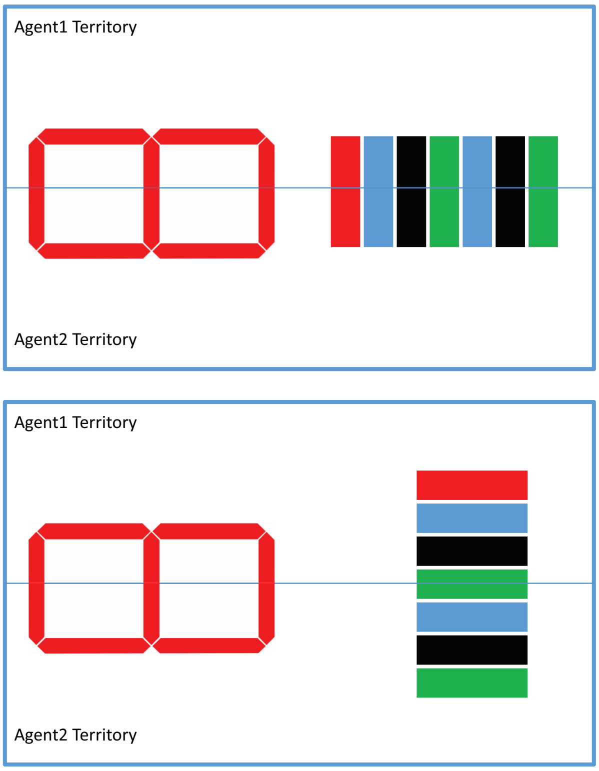

Figure 4.2: Blocks' initial position; the initial position was rotated 90 degrees for every other character in the task meaning if for the current character blocks were initially placed horizontally (top image) for the next they were aligned vertically (lower image).

the first block on the central segment. This is to have both serial (meaning participants take turns in picking up blocks) and parallel actions as it was observed in pilot experiments that some participants performed the task only in a parallel manner (picking up blocks at the same time). In this work, a serial action is considered when one participant waits until his/her partner finishes picking up a block and then moves to pick another block. On the other if both participants moved to pick up blocks together even though with a slight delay in the onset of their motion (not having a synchronous movement) the interaction is considered a parallel action.

There are several works in the literature pointing to the effect of embodiment in the human-robot interaction [121, 122]. Researchers argue that a screen-based or even virtual augmented interaction would lack the embodiment effect [6, 77]. Considering the ultimate goal of the experiment is to collect data to train models for human-robot collaboration, considering the embodiment effect required for the HRC along with the role of physical action in activation of the mirror neuron system, it is considered a physical experiment is more appropriate in this work than a screen-based one.

In two separate experiments for individuals and dyads the following four conditions have been tested:

1. *Distance policy*: Each person has been asked to only pick the closest block.

2. *Colour policy*: Each person has been asked to pick blocks according to an order of colour irrespective of their physical position, i.e., no matter where the block is located the order of colours should be applied first. In terms of the dyadic experiment, each participant is given the same order of colours.

3. *Colour and Distance*: Participants have been asked to follow a colour order and at the same time pick the closest block when there is more than one block of the same colour. The data captured in this condition was used to develop the policy integration layer.

4. *Uninstructed*: In this condition, participants are free to pick blocks without being instructed to follow any order.

These conditions were chosen to create a dataset for a pick-and-place task with a specific order of completion. The order of the task was represented by colour policy. The distance policy was chosen as is suggested that people tend to minimise energy consumption by reaching to the closest object[2, 113, 129]. This was also observed in the uninstructed condition.

4.4 Experimental Design and Participants

A between-subject design was chosen to avoid carry-over effects from one condition to the other. In the first phase of the experiment, 60 individuals took part, 15 people per condition, 44 of which were male and 16 were female. The average age of participants was 28.78 (SD 5.7) ranging from 20 to 48 years old with an average height of 175.4 cm (SD 8.63) ranging from 157 to 191 cm. In terms of handedness, 49 participants were right-handed and 11 were left-handed. All participants reported normal or corrected to normal vision (18 wearing glasses). In the second phase, 96 people in 48 pairs took part, 12 pairs per condition, having an equal number of male and female participants. The average age of participants was 30.92 (SD 10.87) ranging from 18 to 67 years old with an average height of 171.27 cm (SD 10.26) ranging from 148 to 193 cm. For handedness, 85 people were right-handed, 11 were left-handed and 2 reported being dual-handed but used their right hand in the experiment. Participants formed 16 male only, 16 female only and 16 male/female dyads, 37 pairs of both participants being right-handed and 11 pairs of mixed right and left-handed ones. All reported normal or corrected to normal eyesight (44 wearing glasses).

It is noteworthy that despite height differences all participants could reach all objects on the tabletop at all times. To ensure this the height of the participants' chair was adjustable and was equipped free spinning mechanism.

4.5 Neural Field Structure

4.5.1 Structure for Distance policy

For modelling *Distance policy*, the table-top setup of the experiment is mapped into a 2D DNF:

$$\tau\dot{u}(x,y,t) = -u(x,y,t) + h + S(x,y,t) + \int\int w(x-x', y-y')\sigma[u(x',y',t)]dx'dy' \tag{4.1}$$

with

$$w(x,y) = C_{gi} + C_e e^{\frac{-x^2-y^2}{2\sigma_e^2}} - C_i e^{\frac{-x^2-y^2}{2\sigma_i^2}} \tag{4.2}$$

where τ is the time scale, u the activation function over the feature space x (mapped to the length of the table) and y (mapped to the width of the table) at iteration t, h 0 a constant resting level, S an external input or stimulus to the field (with one stimulus for each block and participants' hands). The integral part is to drive lateral interaction in the population with w as interaction kernel and σ a sigmoidal nonlinear threshold function with a scaling parameter β similar to (Eq. 3.21). In the interaction kernel, $w(x,y)$, subscript e stands for the excitatory and i for the inhibitory part of the kernel. By changing values of c_e and c_i the excitatory or inhibitory effects of the kernel can be varied. σ_e and σ_i are to adjust the shape of the interaction kernel (like those presented in Figure 3.4) and C_{gi} decides the amplitude of the global inhibition.

The projected position of the centre point of each block and the wrist position of the participants' wrists is mapped on the x-y plane. The x and y axes are then used as features so that each x-y coordinate of the blocks and wrist is considered the position of an input stimulus to the neural field. Each stimulus is modelled with a 2D Gaussian and the interaction of these stimuli (through lateral interaction shown as integral part of (Eq. 4.1) changes the field activation level in different locations as the input stimuli change due to the agents' motions. The parameters to be learned for this setting are mainly interaction kernel parameters (c_e, σ_e, c_i, σ_i, β, c_{gi}). Having properly trained the model (as detailed in Section 4.6), interaction kernel parameters can change the neural field behaviour such that the response to stimuli will result in activation of the field at the point of interest, respectively the location of the chosen block.

4.5.2 Structure of Colour policy

The *Colour policy* model, is a 1D DNF coupled with a memory trace (Eq. 4.3) having only colours as stimuli:

$$\begin{aligned} \tau_p \dot{P}(x,t)) = \lambda_{build}(-P(x,t) + f(u(x,t)))f(u(x,t)) \\ -\lambda_{decay}P(x,t)(1 - f(u(x,t))) \end{aligned} \tag{4.3}$$

where τ_p is the time scale, $P(x,t)$ is the strength of memory at point x of the DNF with $u(x,t)$ as its activation function and f is a sigmoid function. λ_{build} and λ_{decay} determine the rate of build-up and decay of the memory trace [99]. The training for this structure is like memorising the colour order by demonstrating the order and showing blocks one by one. The memory then forms pre-shapes for the colour order. This model structure is similar to the work by Sandamirskaya and Schöner [100] but implemented in a way that the neural field stays activated to wait in the order until all blocks of the same colour are removed by the participant(s) before moving to the next colour in the order. This means unlike the serial order in[100], there is an additional measure to consider objects of the same colour with the same priority and stop the order to move on to the next colour unless all the previous colour instances are removed. This was simply implemented by an if condition. This is done to simulate tasks with an equal priority of actions in the plan. The parameters of the Colour Policy DNF are chosen to be the same as the ones reported in [100].

4.5.3 Structure of Policy Integrator

This layer of architecture plays an important role in future expansion. Having different policies modelled separately, and integrated through this layer, makes the architecture adaptable to different tasks. For the task at hand, to have a correct prediction on the chosen block, the colour policy model is coupled with the distance policy model. This provides a measure to decide when there exist multiple blocks of the same colour. This means that the colour policy creates a short list of the blocks to be picked and the distance policy model predicts which one will be picked up. This is done by having a DNF similar to the *Distance policy* with only shortlisted blocks as stimuli being implemented and the final block is chosen from the shortlist according to the distance policy. This process occurs naturally in the DNF of the policy integrator as the amplitude of the input stimuli from the output of the colour policy and distance policy models will intensify the neural field activation for the chosen block.

4.5.4 Structure of Negotiation Layer

A simulation of the predicted partners' actions runs simultaneously with the 'own' model in the architecture and their outcome is fed into the negotiation layer. The role of the negotiation layer is to adjust the own decision to the predicted partner's decision accordingly to prevent any conflicts like picking up the same object. This will also adjust the decision based on the plan, so, if the partner's model predicts that the partner would perform the next step, like when a partner reaching quicker to an object, the agent should either move on to the next action or wait for the appropriate moment to perform the next action. This is done by inhibiting own decisions when the model predicts that the partner will perform the same action, or excite own decision when it predicts that the partner is waiting or performing another action. To achieve this, the interaction kernel ($W(x,y)$ in Eq. 4.1) of two DNFs of the own agent model and the partner model, is adjusted based on the human-human interaction experiment. The input to the interaction kernel for the own decision would be the outcome of the predicted decision of the partner and vice versa. In this way, the interaction kernel can either inhibit or excite the outcome of either model depending on the outcome of the decision models. For instance, if the own decision is to pick a specific block this will be input to the interaction kernel of the partner's model, and if the prediction of the partner's decision is to pick the same block this will be input to the interaction kernel of the own decision model. The evolution of these two decisions as two partners are moving towards that specific block, would change the outcome of the interaction kernels so that if one of the agents is approaching the block faster this would lead to inhibition of the other agent's decision and excitation of the faster-approaching agent's decision. This means the desired outcome is achieved by learning when each DNF should be inhibited (activation function being locally or globally deactivated) or excited (activation function either locally or globally being further activated).

4.6 Training Method

The recorded data from individual/dyad participants completing the task was used for training policy models as well as integration and negotiation layers. The training was fairly time-consuming as it took one month to optimise all parameters for the individual phase and as layers added up it became even more computationally expensive. The last training of the negotiation layer for the colour and distance condition took three months.

In the following, the procedure adopted for training are briefly explained, which was applied to all policies (colour, distance, colour and distance) as well as the integration and negotiation layer. The recorded data was split into training and test datasets by randomly choosing the data from two third of the participants for training and the rest for test. With 15 participants for one condition, for example, data from 10 participants was used for training and the rest for testing. All the structures of Section 4.5 were implemented in MATLAB using the DNF toolbox COSINIVA [29]. Parameters of the DNF models were optimised such that it resulted in the desired activation of the field at the correct position and time in the feature space. Basic work on how to train DNFs focuses on gradient-based methods for a local search or evolutionary algorithms for a global search as reported in [65]. A Genetic Algorithm (GA) [81] for a global search was adopted. This was done after other linear optimisation methods like gradient descent were tried and failed to converge to an optimum solution. The recorded data was used as input to the network. Information about the picked block and the predicted decision by the model was used to calculate the error over the whole captured data. To make the results comparable, the same error equation (E_t) for all the policies was used:

$$E_t = \begin{cases} 1 & \text{if} \quad \text{Block}_{p,t} \neq \text{Block}_{n,t} \\ 0 & \text{if} \quad \text{Block}_{p,t} == \text{Block}_{n,t} \end{cases} \tag{4.4}$$

$$E = \frac{\Sigma E_t}{N}, \tag{4.5}$$

where $Block_{p,t}$ is the predicted to-be-next-picked block evaluated at time t by the model and $Block_{n,t}$ is the next picked block evaluated at iteration t and N is the total number of iterations. The overall error E

is then calculated by the average of errors over time. In total, 9 parameters for each field are tuned by the GA: τ, h, c_e, σ_e, c_i, σ_i, β, c_{gi} and σ_w (width of the Gaussian stimuli for the participants' hands). The population size was 200, with stochastic Uniform selection, Scattered crossover and Adaptive Feasible mutation. The training was considered to be finished when the stall generation limit of 50 was reached. The training was performed using pure global coordinates for wrist and blocks. This computation was done in parallel on an HP Z640 14-core machine with 64 GB memory (RAM).

It is worth noting that the training of the negotiation layer was done based on observed human behaviour during joint action. Participants were observed, completing the task, namely taking turns (serial actions) or at the same time (parallel actions) and in few cases a mixture of both serial and parallel actions. When the DNF was optimised for serial actions, the amplitude of global inhibition was larger, while for the parallel actions, it was smaller and the amplitude of local inhibition was larger compared to the serial actions. For a definition of global and local inhibition please refer to Section 3.1.12. Consequently, a training set was formed by mixing data from 3 trials with participants having mainly serial actions and 3 trials with participants having mainly parallel actions.

4.7 Results and Model Validation

Validation of the trained model was performed using a binary performance measure similar to (4.5):

(4.6)
$$P_t = \begin{cases} 0 & \text{if} \quad \text{Block}_{p,t} \neq \text{Block}_{n,t} \\ 1 & \text{if} \quad \text{Block}_{p,t} == \text{Block}_{n,t} \end{cases}$$

(4.7)
$$P = \frac{\Sigma P_t}{N},$$

where $Block_{p,t}$ is the predicted to-be-next-picked block evaluated at time t by the model and $Block_{n,t}$ is the next picked block evaluated at iteration t. The value of P should be close to zero for a trained model well-fitted to the data, assuming that subjects perfectly followed the instructed policy. This measure was used for all the models to have a meaningful comparison.

4.8 Validation Performance

The trained models were tested using separate test sets created randomly from the recorded data of one-third of the participants. The achieved model performance and optimised DNF parameters are reported in Table 4.1, 4.2 and 4.3. As can be seen from Table 4.2, DNF parameters of the negotiation layer when being trained for the Colour policy are the same as for the Distance and Colour condition with serial actions. This indicates that in the Colour only condition, the majority of the participants were taking turns in picking the blocks and serial actions are the main form of interaction. In addition to the DNF activation function and interaction kernel parameters, σ_w representing the width of the Gaussian stimuli for the wrist of participants is presented in this table. While its value has been in the same order for all conditions, for the Distance policy of dyadic experiments the width was found to be much smaller. This is due to requiring higher precision as participants might pick two blocks next to each other. This is why this parameter was also optimised, while the width of Gaussian stimuli for the blocks remained at a constant value of 0.5.

To compute performance measures, the developed models were applied to the recorded data. For individual models, only data from individual experiments was used for training, while for testing data obtained in dyadic experiments was used. In addition, the last row in Table 4.3 shows the performance of the system without a negotiation layer. These numbers were computed by adopting trained individual models for each agent and applying them to all data of the dyadic experiments. The system, in this case, has a relatively low accuracy for colour and colour and distance conditions and shows slightly better performance for the distance condition as each participant picks up the closest block hence reducing the potential conflicts (unless participants of a dyad are of opposite handedness).

Table 4.1: Performance of the policy models for the individual experiment. The accuracy is the mean of accuracy per participant and the standard deviation is calculated by computing the accuracy of the model for each participant and then calculating the standard deviation of these values (the accuracy scale is from 0 to 100%).

Individuals Experiment		Condition			
		Distance	STD	Colour and Distance	STD
Individual	Training set	89.52%	3.72	86.7%	2.94
	Test Set	84.69%	4.28	86.26%	3.08

As an example, snapshots of the activation function of the 2D DNF mapped on the table-top along

Table 4.2: DNF parameters of developed models for individual and dyadic experiments. Parameters of the distance policy and the integrator (colour and distance) are trained from the individual experiment presented in the first two rows. Parameters of the negotiation layer for each condition were trained based on the data from the dyadic experiment presented in the third to sixth rows. The two last rows present two sets of parameters for the negotiation layer of the distance and colour condition for all actions (fifth row) and the last row for trials in which participants were taking turns performing serial actions.

Individuals Experiment	DNF Parameters								
	h	τ	β	c_e	σ_e	c_i	σ_i	c_{gi}	σ_w
Distance	-2.99	19.44	9.52	25.49	1.01	11.02	3.59	-0.54	34.77
Integrator	-2.93	18.94	10.78	25.01	0.88	12.52	3.59	-1.43	36.77
Colour	For parameters see work by Sandamirskaya and Schöner [100]								
Dyads Experiment	DNF Parameters of the Negotiation Layer in each Condition								
Distance	-3.10	17.06	2.28	3.62	16.59	9.89	21.91	-14.94	0.79
Colour	-8.23	17.36	1.55	12.40	16.38	9.12	20.36	-13.34	38.30
Distance & Colour (all actions)	-7.19	13.04	1.29	20.03	10.82	5.35	23.60	-0.22	34.20
Distance & Colour (only serial actions)	-8.23	17.36	1.55	12.40	16.38	9.12	20.36	-13.34	38.30

Table 4.3: Performance of proposed system with and without the negotiation layer in different conditions of the dyadic experiment. No training was done for "Without Negotiation Layer" and the performance result is based on all recorded data. The accuracy is the mean of accuracy for each pair of participants (dyads). The standard deviation is calculated by computing the accuracy of the model based on individual participants and then calculating the standard deviation of these values (the accuracy scale is from 0 to 100%).

Dyads Experiment		Condition					
		Distance	STD	Colour	STD	Distance & Colour	STD
With Negotiation Layer	Training Set	88.46%	5.83	84.64%	2.89	85.31%	2.90
	Test Set	80.57%	2.37	83.58%	1.88	81.39%	3.00
Without Negotiation Layer		72.01%	11.72	57.67%	9.75	65.5%	8.1

with "block layover" are also depicted in Figures 4.3-4.5 to demonstrate how these models work. The small sphere represents a participant's hand and the ellipse is their upper torso position. The lines for the upper and lower arms are drawn approximately as there has been no tracking information for the elbows or shoulders. Figure 4.3 and 4.4, show activation of the neural field and the peak on the approached block, meaning that it is predicted to be picked up by an individual participant. Figure 4.5 is for the same experiment, above showing the DNF activation for participant 1 and below showing the DNF activation for participant 2, respectively. As can be seen, when participant 2 is approaching the blue block, the DNF of participant 1 is inhibited (no activation peak) and the DNF for participant 2 has an activation peak over the blue block, meaning it will be picked up by participant 2.

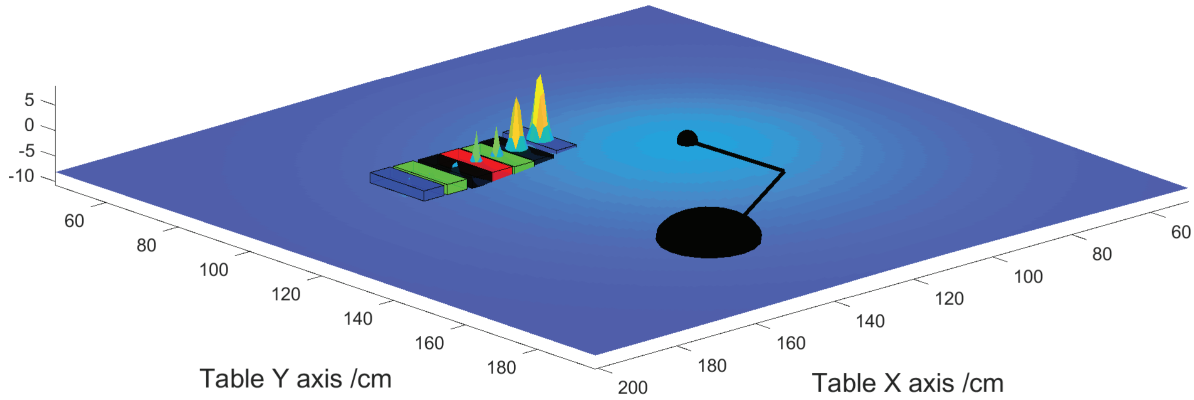

Figure 4.3: Snapshot of 2D DNF activation mapped on the tabletop and overlaid blocks when the participant is approaching the first blue block on the right in an individual trial.

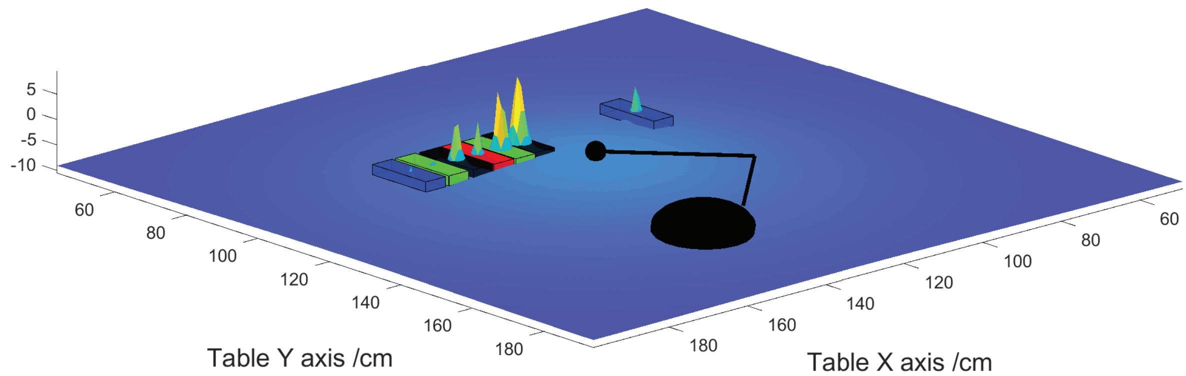

Figure 4.4: Snapshot of 2D DNF activation mapped on the tabletop and overlaid blocks when the participant is approaching the black block after placing the blue block in an individual trial.

4.9 Participants Demographics and Modelling Performance

Considering the physical nature of the human-human interaction in the experiments, ergonomics and the gender of the participants could have an effect on the recorded data and the modelling outcome. To investigate this, the cross-correlation of the training results per dyad participants is evaluated against the following measures, gender difference, meaning the dyad participants were of the same gender or opposite gender, handedness difference, meaning the dyad participants were of the same handedness or the opposite, and hight difference in 4 categories of difference of 5 cm or more, 10 cm or more, 15 cm or more and 20 cm or more, and the age difference of 5 years or more, 10 years or more, 15 years or more, 20 years and more and finally 25 years and more.

No meaningful correlation was found in any of these categories. The model correlation for the

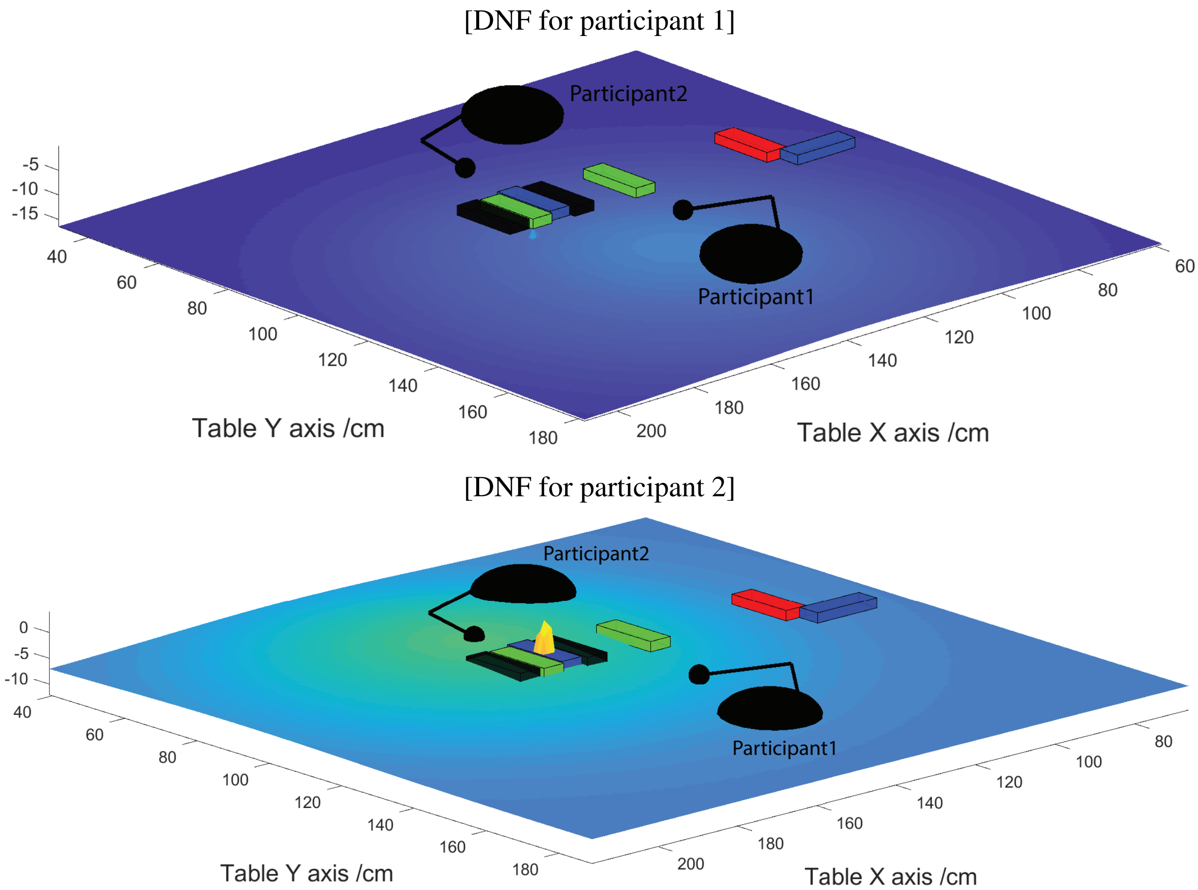

Figure 4.5: Snapshot of 2D DNF activation for participant 1 (above) and participant 2 (below) mapped on the tabletop and overlaid blocks when participant 1 and participant 2 are approaching a blue block at the same time. The neural field of participant 2 is activated with a peak over the blue block predicting this block will be picked up since participant 2 was moving faster than participant 1, while the activation for participant 1 was inhibited.

handedness difference was -0.0.39, for the gender difference, it was -0.2001, and for the height difference of 5 cm or more, 10 cm or more, 15 cm or more, and 20 cm or more, it was 0.0426, -0.1376, -0.1204, and 0.1651, respectively. The correlation for the age difference measures were 0.1736, 0.21396, 0.1534, 0.1076, and 0.0023, for 5 years and more, 10 years and more, 15 years and more, 20 years and more, and 25 years and more, respectively. Nonetheless, considering the number of dyad participants was very limited and not necessarily representative of a normal population this lack of correlation does not necessarily mean no relation. It is, however, crucial to investigate this further in future work with a larger number of participants before drawing any conclusion.

4.10 Comparison to Artificial Neural Networks (ANN)

An all-in-one approach aiming for learning the decision without breaking down the process into policies using a single DFT model failed at the early stages of this work. Thus, it was decided to model policies separately as also indicated in the overall architecture, resulting in the proposed gray-box model. But to also test the DNF and the developed architecture against another black-box technique which has been used in many fields of machine learning including decision-making, further it was decided to compare it to Artificial Neural Networks (ANN). For this purpose, a multi-layer perceptron (MLP) was implemented with h hidden layers of d nodes with *ReLU* activation, *l2-norm* batch normalisation with a regularisation penalty $l2$, dropout at rate d_r, batch size b, using the "Adam" optimiser with early stopping when the accuracy on a 10% validation set had not improved for 20 epochs. To tune these hyper-parameters, a random search in the space of all combinations of hyper-parameters was used as defined by a grid of possible values controlling topology, batch size, normalisation, and dropout. Scikit-learn's Randomised Search method with 5-fold cross validation over 500 iterations was used to tune the meta-parameters, thus in total 100 runs of 500 epochs were used to tune the MLP hyper-parameters, resulting in $h = 4$, $d = 64$, $l2 = 0.01$, $d_r = 0.2$ and $b = 512$.

Unlike the DNF, the MLP model does not have state, so is making an independent prediction at each sampling time-step. However, in practice, the users' decisions only change periodically, and far slower than the observation sampling rate, so access to "memory" could be beneficial. Therefore, also an LSTM recurrent neural network (RNN was implemented), presenting it with a sliding window of samples $(x_{t-w}, x_{t-w+1}, \ldots, x_t)$ from which to predict x_{t+1}. Tuning the meta-parameters, in particular for the batch and window size w proved excessively computationally expensive, so after initial experimentation a topology of two layers of 50 LSTM nodes followed by a single dense layer of 64 nodes, and $d_r = 0.25$ was used, with "data" values of $w = 8$, $b = 512$. The *ReLU* activation functions was chosen as in the past experience they worked well on a range of non-image problems. Some preliminary experimentation with the use of alternative activation functions such as sigmoid and tank was also performed, but the results were not promising.

In training the ANN models a pre-processing stage of the data was required. So that, instead of feeding raw data of (x,y) coordinates of objects used for the DNF models, the data had to be hot-coded with ones for the presence and zeros for absence of objects in the scene for each time sample. In addition, the raw recorded data from the motion capture has a very high sample rate. This had resulted in unnecessarily very large dataset for training DNF models as the training did not need so many coordinates per seconds. Therefore, the data for training the DNF was down-sampled by the factor of 15. However, when the same down-sampled data was used for training of the ANN models it proved to be too small. Hence, the larger recorded datasets had to be used for training of the ANN models.

The results from the two ANN's are shown in Table 4.4 and further compared to the DNF performance in Table 4.4. Please note that the accuracy is similar in most conditions for training data but the ANN model is not comparable for the Colour condition as the DNF model memorises the order and has 100% accuracy in individual experiments.

Table 4.4: The performance of the developed ANN models. The standard deviation values were calculated over 10 runs (accuracy scale is 0 to 100%).

ANN Results			Condition					
			Distance	STD	Colour	STD	Distance & Colour	STD
MLP	Individual	Training Set	79.08%	3.27	86.62%	3.44	92.23%	1.04
		Test Set	43.92%	2.32	77.99%	1.93	85.18%	1.38
	Dyad	Training Set	75.25%	3.19	89.54%	0.76	90.30%	2.05
		Test Set	45.08%	2.14	67.24%	2.31	73.69%	2.64
RNN	Individual	Training Set	61.82%	1.18	78.67%	1.23	79.71%	0.71
		Test Set	24.72%	1.35	64.32%	0.81	63.68%	1.02
	Dyad	Training Set	59.21%	5.26	81.84%	0.34	85.89%	0.82
		Test Set	18.76%	2.75	67.81%	1.39	46.61%	0.88

As can be seen, despite having "memory", the recurrent network did not always achieve the same level of training accuracy as the more basic MLP, and the gap between accuracy on the training and test sets was typically far greater - a classic sign of "overfitting". The MLP also displays some signs of overfitting and never reaches the test accuracy observed with the DNF. Possible reasons for this are as follows:

- The form of the model: the MLP does not have state, and so cannot take advantage of the differences between the rate of decision-making and sampling. This should have been ameliorated by the use of LSTM nodes in the early layers of the recurrent model - and was for all but the individual-distance combination.

Table 4.5: Comparing performance of DNF, MLP and RNN in different conditions of the individual and dyadic experiments. Bold type indicates highest value in each row, but it is not intended to assert statistical significance (accuracy scale is 0 to 100%).

Condition		Accuracy (%)		
		DNF	MLP	RNN
Distance				
Individual	Training set	**89.52**	79.08	61.82
	Test Set	**84.69**	43.92	24.72
Dyad	Training set	**88.46**	75.25	59.21
	Test Set	**80.57**	45.08	18.76
Colour				
Individual	Training set	**100**	86.62	78.67
	Test Set	**100**	77.99	64.32
Dyad	Training set	84.64	**89.54**	81.84
	Test Set	**83.58**	67.24	67.81
Colour & Distance				
Individual	Training set	86.7	**92.23**	79.71
	Test Set	**86.26**	85.18	63.68
Dyad	Training set	85.31	**90.30**	85.89
	Test Set	**81.39**	73.69	46.61

- Insufficient computational budget for meta-parameter tuning and training: direct comparisons are difficult as the MLP and RNN were built on a 48-core processor exploiting two fast GPU cards. However each was given a week's runtime, which approximately equates to the month on a slower machine for the DNF.

- The greater complexity (number of weights to learn) is greater than the DNF and consequently requires more training data to avoid overfitting. This should have been ameliorated by the use of early stopping. Nevertheless, both MLP and RNN have several categorical meta-parameters to tune, followed by thousands of continuous-valued weights for which values must be optimised.

- The algorithm used to optimise the model parameters: both used stochastic algorithms, but the DNF parameter space was searched using the *global* search of an evolutionary algorithm, whereas because of the far greater search space size, the MLP and RNN networks used a variant of *local* search (gradient descent).

In summary, the evidence suggests that the MLP is outperformed by the DNF due to its lack of memory. For some tasks, the RNN approached or bettered, the accuracies of the DNF in its performance on the training data. However, to an even greater extent than the MLP, it suffered from "overfitting", so that predictive accuracy on unseen test data was poor. Given greater computational budget for tuning the meta-parameters (such as network size, "early stopping" rule) could possibly have been improved. However, the greater simplicity of the DNF makes global optimisation feasible, and reduces the likelihood of overfitting, making it faster and simpler to tune to a good level of accuracy for different tasks.

4.11 Discussion

The proposed decision-making structure based on DNFs is developed and tested comprehensively for a simple task. However, the main primitive underlying action of this task, pick-and-place, is indeed part of many more complex tasks people perform in their day-to-day life. Hence, we strongly believe that it is possible to apply this structure also to more complicated tasks, with one example being presented in Chapter 6. There might be a need to integrate more policy models though and to eventually also update them over time. This modularity, however, is considered a clear advantage of the proposed architecture as it allows evolution over time by adding or removing required policies for different tasks with different degrees of complexity.

On the other hand, as pick-and-place is a natural part of many collaborative tasks, we consider it reasonable to re-use the already trained *Distance policy* without any need for further training (as is done in the experiment in Chapter 5) and to extend the idea of the *Colour policy* to any preferred order of actions. Hence, there would be only a need to train further involved policies as well as the integration layer. We further assume the negotiation layer would not need to be re-trained as long as the main nature of the task remains a pick-and-place process, each agent is performing an action either in series or in parallel to the other agent to achieve the shared goal and there is no shared action like carrying a large object together as well as the negotiation layer has been trained with a sufficiently rich dataset to capture a large series of eventual cultural or personal preferences. Otherwise, offline re-training or better online updating of the negotiation layer may be required. All these assumptions, however, need to be still proven in future work.

An all-in-one approach aiming for learning the decision without breaking the process down into policies with a single model based on a DNF failed at the early stages of this research, hence, it was hypothesised a way forward is to model policies separately as also indicated in the overall architecture and resulting in the proposed grey-box model. A clear benefit of modelling each policy separately and integrating them in a second step rather than a learning-all-in-one approach is that the properties of individual policies can be considered. For instance, in the presented task in this chapter, the *Colour policy* is learnt by memorising the colour order and the *Distance policy* model is trained based on the recorded data optimising DNF parameters using a GA. Using the integration layer makes it possible to have several policy models trained with various approaches and integrated into the overall architecture. Furthermore, having the policy models developed separately makes it possible to easily add more policies to the architecture depending on future tasks. Some possible extensions are explained in the last chapter as future work.

In addition, the chosen task and instructions, make it possible to generalise the developed models for a complex task without requiring a complete retraining of the models. As mentioned before, a *Distance policy* is an integrated part of most pick-and-place tasks. The colours order also has been chosen in a way so that *Colour policy* could be generalised for other tasks with orders of actions, either with serial order like picking the red block first then the blue one, or having a parallel order for two actions with the same priority, like picking either of the blue blocks. This is clearly demonstrated in the experiment results of Chapter 6.

When training the models, despite the close performance of the RNN models to DNF models for the training data in some conditions, for the test data, DNF outperformed both RNN and MLP. It can be due to the nature of DNF models and having a travelling wave for the movement of the blocks and participants, while MLP does not have state, and so cannot take advantage of the differences between the rate of decision-making and sampling. This should have been ameliorated by the use of LSTM nodes in the early layers of the recurrent model. However, insufficient computational budget for meta-parameter tuning and training of RNN models considering their greater complexity (number of weights to learn) than the DNF model consequently requires more training data to avoid overfitting.

It is noteworthy that the GA itself is governed by several parameters and operator choices, most notably whether it was allowed to re-evaluate duplicates (This was not restricted). Moreover, the GA is evolving the weights for the DNF, whereas the random search is tuning the MLP hyper-parameters, but the MLP weights are being tuned by a sophisticated meta-heuristic (Adam). Therefore, it could be argued that although we have attempted to achieve parity of computational effort within the use of standard toolkits (to replicate research), it is probably never possible to guarantee exact equality.

As the model was developed using a supervised learning approach, the outcome was compared to ANN as another supervised learning method. Nonetheless, an attempt was also made to apply POMDP on the datasets. Wang [125] found, although initial implementations using POMDP based on an artificially created dataset seemed promising when applied to the real dataset, the POMDP-based model completely failed. This is likely due to the noise in the data. The recorded data consists of tracking coordinate frames of participants' motion and there are many short temporary losses of tracking. This, however, does not affect the DNF-based model as its dynamic nature damps a sudden change in the input stimuli. In addition, when lacking the probability of events and consequences, Reinforcement-Learning(RL)-based approaches typically require many iterations so would be less practical for human-machine collaboration.

In the human-human interaction experiments, the majority of people tended to minimise their energy consumption by picking the closest object. This has been also observed when participants were asked to perform the task without any instruction, making *Distance policy* the main naturally chosen policy in a pick and place task. In this case, using the system only with *Distance policy* models resulted in 87.78% prediction accuracy.

The negotiation layer may also facilitate safe human-robot collaboration. The collaboration will be safer as the Negotiation Layer reduces the chances of conflicts and, thereby, unwanted or unintentional contact since the human's action is directly affecting the robot's decisions. As Table 4.3 presents, the results showed that in dyadic scenarios, incorporating the negotiation layer improved the model performance, for example, the highest improvement was for the *Colour Policy* experiment by 26%.

Using DNFs along with a global optimisation approach like a GA to optimise its parameters has one disadvantage, i.e., it is computationally expensive. However, one could expect this to be mitigated in the future by emerging faster high-performance computers. For example, in this work during the training phase, a machine with 14 cores was used to run the optimisation in parallel making the training significantly faster than using a normal computer. Gradient descent approaches (available in Matlab) were also tested for optimising the parameters, however, they did not converge and ended with a relatively high error. This suggests that GA was a good choice for optimising DNF parameters as it can escape local minima. On the other hand, using a global search approach has made the training highly computationally expensive. Although this could be alleviated by advances in computing power and parallel processing, at present, it has limited the training phase. At the same time, considering the dynamic behaviour of the DNF, it is highly resilient to variation in its input and, in most previous work (cited in Chapter 3.1.12) for less complex systems, its parameters were chosen manually through expert tuning.

Another advantage of using DNF for modelling decision-making is that it can capture the dynamic process of human decision-making. For example, in the data from the human-human interaction pick-and-place task, any changes in the human motion indicating a change in the decision would dynamically change DNF activation on-the-fly and the predicted decision would be updated. When using a 2D DNF this process can also be visualised for a better understanding of the decision-making process and evaluation of the model performance.

CHAPTER 5

Human-Robot Collaboration Experiments

This chapter is aimed for evaluating the performance and perceived naturalness of the developed architecture in a human-robot collaboration task. For this, the architecture was implemented on a robotic arm and human participants were asked to collaborate with the robot in performing exactly the same task as that which had been used for data collection and training the models. A Wizard of Oz test was used in which a human teleoperated the robot as a baseline and compared it to the performance of the architecture with and without the negotiation layer.

In addition to objective measures, participants' answers to the Godspeed questionnaire [7] taken at the end of each condition as well as the obtained answers to three post-experimental interview questions were evaluated. Results reported in Section 5.8 and discussed in Section 5.9 indicate that the introduced architecture with the negotiation layer achieves similar performance as the baseline, while the architecture without the negotiation layer performs statistically significantly worse.

5.1 Experimental Design

To evaluate the developed decision-making model, three conditions are tested in a human-robot collaboration experiment: First, the complete model that includes the negotiation layer, second, the model without the negotiation layer, and finally, the human ('Wizard of Oz') decision-maker.

In pilot experiments where a complete within-subject design was tried, it was noticed that the robot behaviour in each condition strongly affected how participants interacted with the robot in the subsequent conditions, for example, when there were conflicts in the "no-negotiation" condition people changed the way they worked with the robot to avoid those conflicts and adapted to the robot, for example by trying to take turns rather than collaborating.

Thus, to overcome this problem and to reduce the total number of subjects required for a full between-subject design, it was decided to run two within-subject experiments comparing the baseline (Wizard of Oz) to the model with a negotiation layer and another one for comparing the performance of the model with and without the negotiation layer. This way it was possible to reduce the required number of subjects by one-third, which helped to conclude the experiment in a timely manner due to increased difficulty in conducting group behavior experiments during the coronavirus pandemic.

Ethical approval for these experiments was obtained from the ethics committee of the University of the West of England (reference number: UREC16-17.03.10).

5.2 Experimental Setup

The experimental task was chosen to be the same as that used during the data collection phase for training the decision-making models and consists of a table-top pick-and-place task which is explained in Section 5.4. The main difference consisted in the fact that participants were asked to collaborate with a Franka Emika Panda robotic arm instead of a human collaborator.

The experimental setup is shown in Figure 5.1. The dominant hand of the participant was equipped with reflective balls to allow its tracking by a Vicon motion capture system. The robot was controlled in real-time using the libfranka library in C++. The decision-making model was implemented in MATLAB and passed its final decision on for execution to the robot controller via TCP/IP socket communication. The implementation foresaw that the robot controller was not continuously receiving the decision from Matlab, but only at specific times in both conditions (Wizard or Model decision maker). Hence, once the robot initiated an action, the robot could not change its behaviour anymore but had to conclude this action.

Participants repeated the same task in two conditions. In the first experiment, the performance of the decision-making model with and without the negotiation layer were compared. In the second experiment, the model with the negotiation layer was compared to the baseline (Wizard of Oz), where a human decision-maker commanded the robot. The order of the two conditions along with the blocks' locations (horizontal vs vertical alignment) was randomised. The robot controller and motion planner remained the same in all conditions.

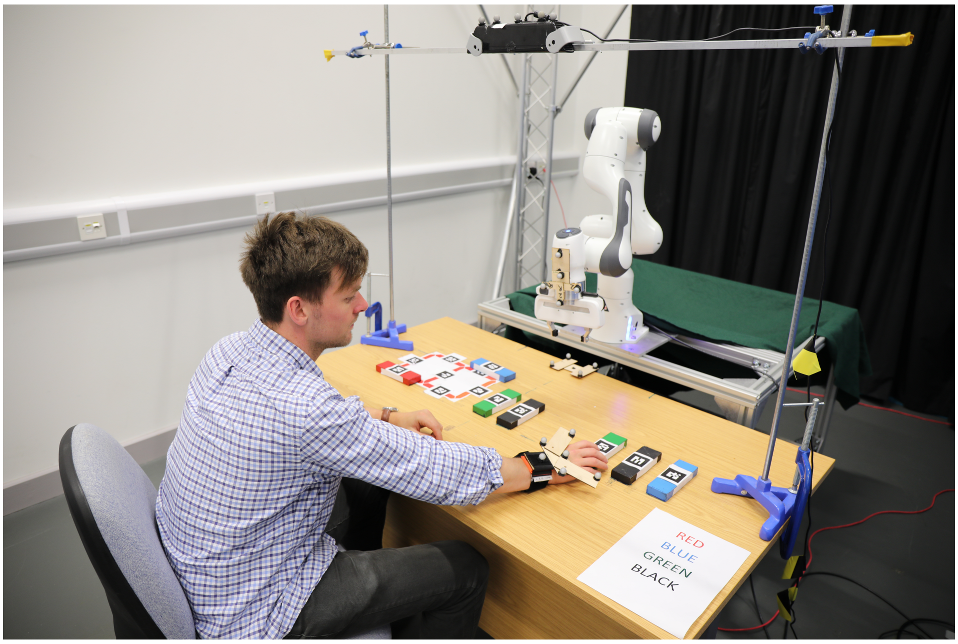

Figure 5.1: Experimental Setup.

5.3 Wizard Protocol

In the baseline condition, the wizard sees the scene from the MS Kinect camera mounted over the table to also track the blocks marked by AR markers. To guarantee a consistent behaviour of the human decision-maker, the following protocol was given to the wizard to follow:

The wizard makes the first decision as soon as the robot is in the "ready position". This is done by entering a number between 1 to 7, 1 for the far-right and 7 for the far-left block from the robot perspective, or entering 8 for just waiting. Later the wizard needs to confirm this decision when the robot goes to the "ready-to-grasp" position. As soon as the robot reaches the "ready-to-grasp" position, the wizard must make the final decision within one second or as soon as the human participant makes his/her move (whichever is faster). At this stage, the wizard has 3 options, either to confirm the decision and continue picking the same block, change to another block, or just wait. If the decision is waiting, the wizard needs to make another decision again within one second until a block is chosen. This will continue until the task is complete. It is noteworthy to mention that the wizard needs to make a decision even if the decision is waiting so that the robot control loop is reset and the robot would be ready for the next action. If no decision is made by the wizard, the robot would stop working.

5.4 Task

The experimental task was kept the same as used during the data collection phase. Participants were asked to use provided coloured blocks to form some alpha/numeric characters on a 7-segment template, according to a provided colour order. Each participant was asked to pick blocks one by one, however, they were just told to "work with the robot" and were free to either take turns or pick a block at the same time as the robot. The alpha/numeric characters were "H, 2, 5, 3, E, 6, 9, 8". After completing each character, the experimenter was resetting the blocks and randomised their position. The order of the character was also randomised and the experimenter told the participants which character to assemble next every time the scene is reset.

5.5 Participants

5.5.1 First Experiment (model with negotiation layer vs model without negotiation layer)

In total 40 participants (22 male) took part in the experiment. Participants were staff and student members of the University and had an average age of 31.9 (STD = 8.47) ranging from 21 to 54 years with an average height of 173.22 cm (ST D = 8.98) ranging from 158 to 191 cm. 38 participants were right-handed, and all reported normal or corrected-to-normal eyesight (14 wearing glasses).

5.5.2 Second Experiment (model with negotiation layer vs human decision-maker)

Again, in total 40 new participants (29 male) took part in the experiment. Participants were staff and student members of the University and had an average age of 31.8 (STD = 9.5) ranging from 20 to 61 years with an average height of 175.05 cm (STD = 9.35) ranging from 157 to 200 cm. 37 participants were right-handed, and all reported normal or corrected-to-normal eyesight (18 wearing glasses).

5.6 Subjective and Objective Measures

In terms of objective measures, the task completion time, the robot task share, and the number of conflicts were recorded. The robot task share is calculated as the mean value of the number of blocks the robot picks divided by the total number of blocks required for the character being created. A conflict is considered when the robot tries to pick the block that the participant has just picked up or is about to be picked up.

As for the subjective measures, at the end of each condition, participants were asked to answer the Godspeed questionnaire. As the questionnaire is designed for use with humanoid robots, they were asked to only focus on the robot's behaviour when choosing and picking the blocks. Finally, at the end of the experiment, participants were additionally asked the following three questions:

- **Q1.** What do you think was the difference between the first and second conditions?
- **Q2.** Which condition did you prefer?
- **Q3.** Which condition was more like working with a human partner?

5.7 Data Analysis

To evaluate the model, a statistical analysis of the recorded objective and subjective measures was performed. Since data was found to be not normally distributed for some of the objective data, an Analysis of Variance (ANOVA) was chosen instead of a classical t-test, which would be unreliable, while there is a strong body of research suggesting the ANOVA F-test to be not sensitive to normality assumptions [18, 51, 61, 78]. In addition, using ANOVA allows the Tukey test to be performed for Post-hoc analysis. This allows for a robust comparison that ensures a significant difference exists even for marginal situations. By contrast, the t-test would fail in these situations.

Considering that the absence of a significant difference in the Null hypothesis test does not necessarily imply a similarity, a further Bayesian comparison [12] of the conditions of the second experiment was performed to establish similarity.

In addition, as the data for the number of conflicts is rather sparse, a difference of proportion hypothesis test (DPHT) was performed for this measure in the second experiment. To do so, the total number of instances in which there was "no conflict", "one conflict", and "two conflicts" for both the model and human decision-maker was derived from the data and the z score of the DPHT was calculated according to

$$z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1-\hat{P})(\frac{1}{n_1} + \frac{1}{n_2})}} \tag{5.1}$$

where n_1 and n_2 are the number of samples for each condition, here both equal 40. Both, $\hat{P}_1$ and $\hat{P}_2$ are found by dividing the number of conflicts by the number of samples for each condition. Finally, $\hat{P}$ is found with

$$\hat{P} = \frac{n_1\hat{P}_1 + n_2\hat{P}_2}{n_1 + n_2} \tag{5.2}$$

Having the z score, the p-value can be found using an inverse normal distribution function like $NORM.S.INV$ in any statistical package.

As for qualitative analysis, NVivo was used to extract keywords used by participants for describing each condition. A Word Frequency Query and a Matrix Coding Query were performed. The former is to analyse the frequency of the words with the same stem that participants used to describe each condition and the latter is to analyse the attitude (Positive, Negative, Mixed, Neutral) of the participants towards each condition.

5.8 Results

5.8.1 The First Experiment (model with negotiation layer vs model without negotiation layer)

The data distribution was analyzed using the Jarque-Bera method. Not all the data had a normal distribution as shown in Table 5.1, however, as previously mentioned the ANOVA F-test is still being considered a reliable approach for a null hypothesis test.

Table 5.1: Jarque-Bera Normal Distribution test and Cohen d effect size for the objective measures. H=0 means data has a normal distribution.

Objective Measures	Condition	JB test H	JB test p	Cohen D Effect Size
Robot Task Share	With negotiation	1	0.0268	1.0194
	Without negotiation	0	0.1935	
Task Completion Time	With negotiation	0	0.1718	0.4234
	Without negotiation	1	0.0219	
Conflicts	With negotiation	0	0.0647	2.3964
	Without negotiation	0	0.1860	

A one-way analysis of variance (ANOVA) followed by a Tukey test were performed on the recorded objective and subjective measures. All objective measures for the condition with the negotiation layer were found to significantly outperform the condition without such layer. The mean task completion time with the negotiation layer was significantly ($F = 17.03, p\ \ 9.15e-05$) less than without the negotiation layer. Nonetheless, 10 participants mentioned the difference between the two conditions was that the robot was "faster" or "quicker" without the negotiation layer. The mean number of conflicts was also significantly ($F = 205.82, p\ \ 1.4e-23$) lower in the condition with the negotiation layer compared to the condition without such a layer. This was further confirmed in qualitative interview data by participants describing the robot's behaviour without the negotiation layer as making more "mistakes" (13 participants), being more "aggressive" or "irritating" (4 participants,) more "competitive" (2 participants) and using other adjectives like "foolish/dumber", "intrusive", "ignored" and "unpredictable". The mean robot task share was found to be significantly ($F = 33.9, p\ \ 1.23e-7$) higher in the condition with the negotiation layer. Results of the statistical tests performed on the objective measures are reported in Appendix A.1.

Table 5.2: Mean value and standard deviation of the objective measures.

Objective measures	Conditions	
	With negotiation	Without negotiation
Mean Completions Time (s)	440.2 (STD=64.18)	495.8 (STD=57.78)
Mean Robot Task Share	0.451 (STD=0.03)	0.416 (STD=0.021)
Mean number of Conflicts	0.73 (STD= 0.78)	5.85 (STD=2.28)

As for the subjective measures, although all the scores for the model with negotiation layer were found to be higher, only data in two categories of the Godspeed questionnaire, namely "Anthropomorphism" ($F = 8.4, p\ \ 0.0049$) and "Perceived Intelligence" ($F = 20.83, p\ \ 1.84e-5$) scored significantly better when using the model with negotiation layer. However, no significant difference was found for the other categories of "Animacy", "Likeability", and "Perceived Safety", see also Figure 5.2 reported in Appendix A.1.

Table 5.3: Jarque-Bera normal distribution test and Cohen d effect size for the objective measure. H=0 mean data has a normal distribution.

Measure	Condition	JB test H	JB test p	Cohen d effect size
Anthropomorphism	With Negotiation	0	0.3655	0.6694
	Without negotiation	0	0.3312	
Animacy	With negotiation	0	0.5	0.4757
	Without negotiation	0	0.5	
Likeability	With negotiation	0	0.5	0.5016
	Without negotiation	0	0.5	
Perceived Intelligence	With negotiation	0	0.2866	0.9696
	Without negotiation	0	0.5	
Perceived Safety	With negotiation	0	0.0538	0.4234
	Without negotiation	0	0.2256	

Table 5.4: Mean value and standard deviation of the subjective measures of Godspeed questionnaire.

Subjective measures	Conditions	
	With negotiation	Without negotiation
Anthropomorphism	2.815 (STD=0.86)	2.315 (STD=0.68)
Animacy	2.95(STD=0.77)	2.66 (STD=0.77)
Likeability	3.565 (STD= 0.585)	3.275 (STD=0.72)
Perceived Intelligence	3.515 (STD=0.54)	2.835 (STD=0.77)
Perceived Safety	3.775 (STD= 0.63)	3.6 (STD=0.646)

A further analysis was performed on the data when sorted, based on the order of trial rather than conditions and no significant difference was found in the data based on trials' order.

In addition, a word frequency analysis of answers to the first question of the post-experiment interview was performed using NVivo software by looking for words with the same stem and more than 5 letters, see tables A.17 and A.18 in Appendix A.2. The NVivo word maps created from the word frequency query are shown in Figures 5.3 and 5.4.

Since NVivo could not separate words based on the way they are used in the sentence, like using an adjective positively or negatively by adding "not" before it, a further Matrix Coding Query was

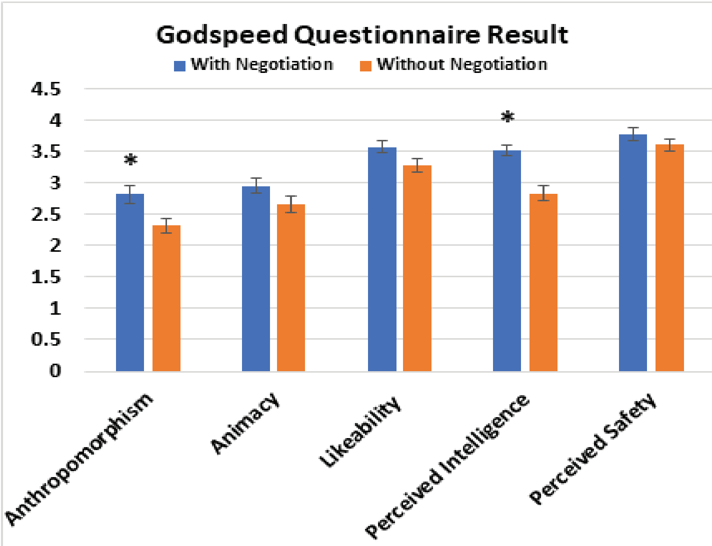

Figure 5.2: Bar graph for Godspeed questionnaire. Mean values of all categories for the first experiment. Error bars are $\pm 1 SEM$ (Standard Error of Mean)

Figure 5.3: NVivo word map for the answers to the first interview question for the condition without negotiation layer. Note, the positive adjectives like responsive and pleasant were used in negative form with "not" to describe the robot in this condition but the software cannot show that.

performed, see results reported in table 5.5. Attitudes are coded manually in such a way that if there are only positive comments about the robot it is coded as "Positive", for example, a participant describes the robot running the model with the negotiation layer as "it was like working with another person. It was

Figure 5.4: NVivo word map for the answers to the first interview question for the condition with negotiation layer.

like having an interaction." or if only negative comments are used it is coded as "Negative" like this statement of the same participant describing the robot using the model without negotiation layer as "It felt like working with a machine and the robot was just doing its job". If mixed positive and negative adjectives are used it is coded as "Mixed" such as "it was slower but easier to work with" as a description of the robot with the model with negotiation layer. If there are neither positive nor negative comments it is coded as "Neutral".

Table 5.5: NVivo Matrix Coding Query for conditions vs attitudes. Numbers show the number of times participants have described a condition with each attitude.

Matrix Coding Query		Attitude			
		A : Mixed	B : Negative	C : Neutral	D : Positive
Condition	With negotiation	6	2	0	23
	Without negotiation	7	24	0	3

Answering the second interview question, 30 participants mentioned they preferred the condition with the negotiation layer. The other 10 participants who preferred the *without negotiation* condition mentioned the robot was faster and that's why they preferred it. This is despite the actual robot movement being the same in both conditions, however, one of the effects of the negotiation layer on the robot behaviour was to wait and give way to the participants.

Answering the last question of the interview, 33 participants said the robot with the negotiation layer was more human-like and one found no difference. 4 Participants mentioned the robot without the negotiation layer was faster hence they found it more human-like and one described the robot's behaviour as "unpredictable" and hence more human-like.

5.8.2 Second Experiment (model with negotiation layer vs human decision-maker)

Data captured in this experiment was tested for normality distribution using the Jarque-Bera test. Not all data was found to have a normal distribution as indicated in table 5.6.

Table 5.6: Jarque-Bera normal distribution test and Cohen d effect size for the objective measures. H=0 means data has a normal distribution.

Objective Measures	Condition	JB test H	JB test p	Cohen d Effect size
Robot Task Share	With negotiation	0	0.5	0.1388
	Human decision-maker	0	0.3137	
Task Completion Time	With negotiation	0	0.2592	0.4270
	Human decision-maker	0	0.5	
Conflicts	With negotiation	1	0.0068	0.2195
	Human decision-maker	1	1.0000e-03	

A one-way ANOVA was performed to analyse the data captured in this experiment. Unlike the first experiment, no significant difference was found for any of the objective or subjective measures. Nonetheless, the human decision-maker was found to descriptively outperform the model, even though marginally, in most categories. In terms of the objective measures, data were normally distributed for Robot Task Share and Task Completion Time but not for the number of Conflicts in either condition. The result of the Jarque-Bera test and Cohen d effect size are shown in table 5.6.

Table 5.7: Mean value and standard deviation of the objective measures.

Objective measures	Conditions	
	With negotiation	Human decision-maker
Mean Completions Time (s)	486.3 (STD=78.75)	463.95 (STD=75.85)
Mean Robot Task Share	0.463 (STD=0.038)	0.469 (STD=0.032)
Mean number of Conflicts	0.55 (STD= 0.75)	0.45 (STD=0.63)

Table 5.8: Jarque-Bera Normal Distribution test and Cohen D Effect Size for the subjective measure. H=0 mean data has a normal distribution.

Measure	Condition	JB test H	JB test p	Cohen d effect size
Anthropomorphism	With negotiation	0	0.5	0.2224
	Human decision-maker	0	0.5	
Animacy	With negotiation	0	0.5	0.3044
	Human decision-maker	0	0.5	
Likeability	With negotiation	0	0.5	0.1151
	Human decision-maker	0	0.1033	
Perceived Intelligence	With negotiation	0	0.0512	0.1787
	Human decision-maker	0	0.5	
Perceived Safety	With negotiation	0	0.4711	0.0824
	Human decision-maker	0	0.5	

As for subjective measures, similarly, there was no significant difference found for any categories of the Godspeed questionnaire, although the human decision-maker scored descriptively better in all. The mean values of these categories are depicted in Figure 5.5. All the data were normally distributed and the results of the Jarque-Bera test and the Cohen d effect size are reported in table 5.8.

Table 5.9: Mean value and standard deviation of the subjective measures derived from Godspeed questionnaire.

Subjective measures	Conditions	
	With negotiation	Human decision-maker
Anthropomorphism	2.875 (STD=0.665)	3.025 (STD=0.71)
Animacy	2.99 (STD=0.73)	3.196 (STD=0.71)
Likeability	3.615 (STD= 0.73)	3.68 (STD=0.67)
Perceived Intelligence	3.69 (STD=0.73)	3.805 (STD=0.59)
Perceived Safety	3.683 (STD= 0.75)	3.73 (STD=0.77)

Having no significant difference in the Null hypothesis test, a further Bayesian comparison [12] of the conditions was performed to test for similarity. To do so, all the data was normalised based on the maximum value of each category. The results of this comparison between the model with the negotiation layer and human decision-maker for all the objective and subjective measures are presented in Tables 5.10 and 5.11. In these tables, the probability of similarity, and the probabilities of either condition outperforming the other are depicted. The last column in the tables shows the width of the Region of Practical Equivalence (ROPE). ROPE is in the same order as the data; since the data is

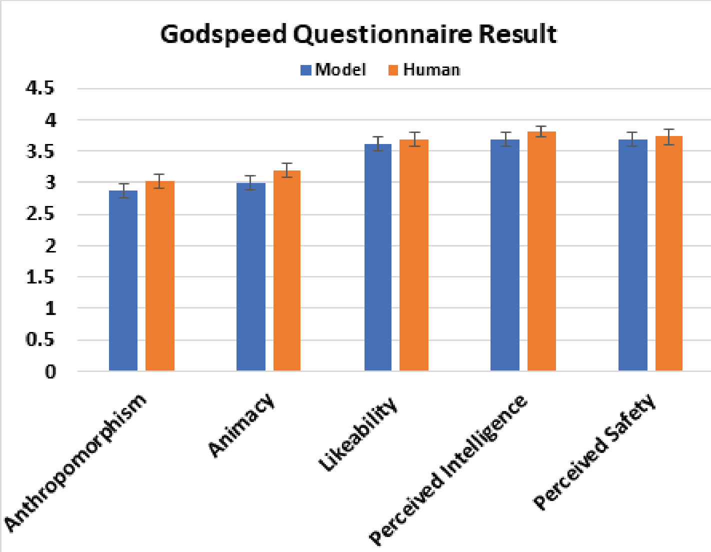

Figure 5.5: Bar graph for the Godspeed questionnaire. Mean value of all categories for the second experiment. Error bars are $\pm 1 SEM$ (Standard Error of Mean)

normalised, the ROPE value was chosen between 0 and 1. Having a high probability of similarity with a smaller ROPE is desirable as it means even a higher likelihood of similarity than with a larger ROPE closer to 1.

Table 5.10: Bayesian comparison of the model and human decision-maker for objective measures.

Objective Measures	Model is better	Difference is negligible	Human is better	ROPE
Robot Task Share	1.46e-14	0.9999999999	4.09e-13	0.1
Task Completion Time	3.9e-08	0.99999996	4.32e-13	0.1
Conflicts	0.00088	0.999	2.34e-06	0.5

Table 5.11: Bayesian comparison of the model and human decision-maker for subjective measures.

Measure	Model is better	Difference is negligible	Human is better	ROPE
Anthropomorphism	5.89e-05	0.9867	0.0133	0.1
Animacy	2.195e-05	0.9671	0.0328	0.1
Likeability	3.52e-05	0.9992	0.00073	0.1
Perceived Intelligence	6.50e-05	0.9943	0.0057	0.1
Perceived Safety	0.000125	0.99883	0.00104	0.1

Having rather sparse data for the number of conflicts, a difference of proportion hypothesis test (DPHT) was performed for this measure. The total number of instances in which there was "no conflict", "one conflict", and "two conflicts" for both the model and human decision-maker was derived from the data as shown in table 5.12. The z value of the DPHT was calculated according to eq 5.1. Determined z values are shown in table5.13. Values indicate no significant difference for a significance level $\alpha = 0.05$ when comparing the model with decision layer and the human decision-maker.

Table 5.12: Number of conflicts per condition per occurrence.

number of conflicts	Conditions	
	With negotiation	Human decision-maker
no conflicts	24	25
one conflict	10	12
two conflicts	6	3

Table 5.13: The z and p values of the Difference of Proportion test for the number of conflicts.

Number of Conflicts	z	p
no conflicts	-0.2295	0.591
one conflict	-0.5008	0.692
two conflicts	1.0615	0.144

Similar to the first experiment, data was sorted based on the order of trials, however, no significant difference was found between the first and second trials.

Using NVivo, a word frequency analysis of the answers to the first question of the post-experiment interview was performed by looking for the words with the same stem and more than 4 letters. Results are reported in tables A.19 and A.20 of Appendix A.2. The NVivo word maps created from the word frequency query are shown in Figures 5.6 and 5.7.The first obvious observation is that when describing the robot with the model with the negotiation layer, participants used different words like it was "waiting" (21 times) for me, "hesitant" (3 times), or "slower" (4 times) or "calmer" (1 time) showing they perceived the robot was not as fast as the human decision-maker. While for the human decision-maker condition, words like "faster" (15 times), "quicker" (6 times), "snappier", "wilder", and "determined" (each of which 1 time) were used.

Another observation is the use of the words "difference" and "similar", which were used by participants as they did not find any difference between the two conditions. This is later categorised as a neutral attitude in the NVivo Matrix Coding query.

Analysing the results of the Word Frequency Query shows that the robot implementing the model

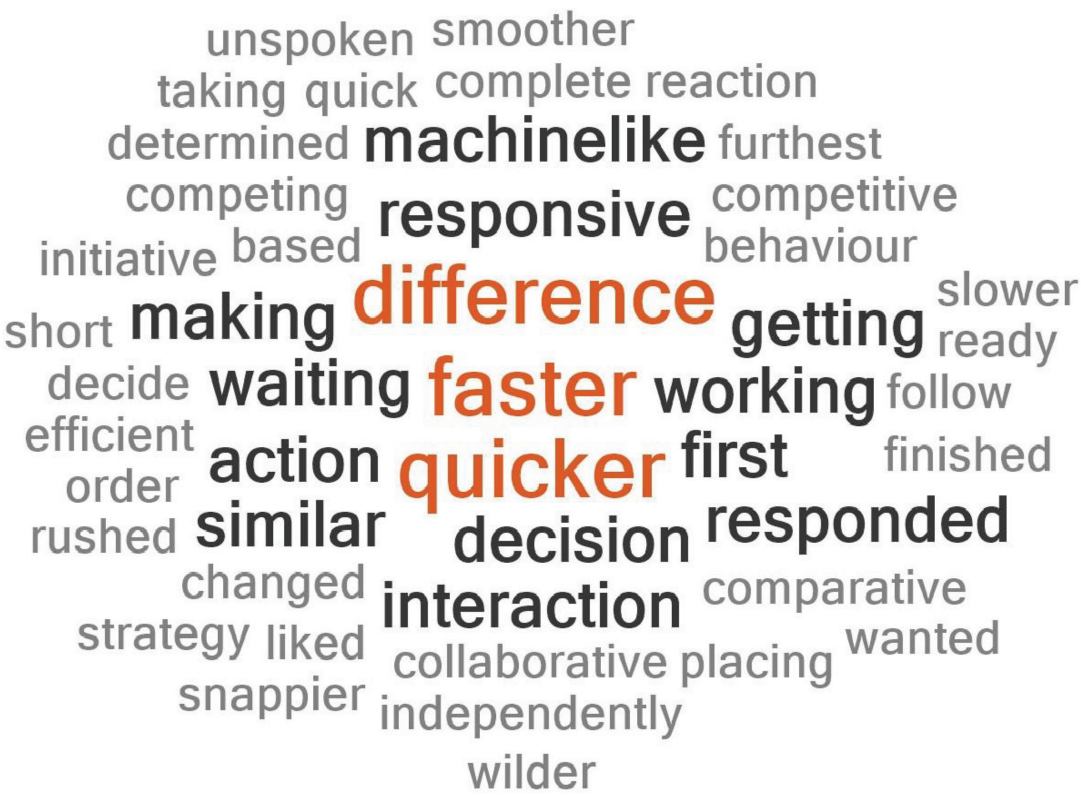

Figure 5.6: NVivo Word Frequency Map for the answers to the first interview question for the conditions human decision maker (robot controlled by a human decision-maker).

with negotiation layer to be described positively using words like "changed behaviour/decision/action" (5), "responsive" (4), "working together" (3), "interactive" (2), "human-like" (2), "polite" (2), "collaborative/cooperative" (2), "considered/considerate" (2), "accommodating" (1), "calmer (1), whereby some were used in completely positive statements and some in a mixed one. For instance, a participant described the robot implementing the model with the negotiation layer as "It was a bit slower and waits for me to pick first. Less aggressive and changed its behaviour if approaching the same block.", in a mixed statement if being slow is considered to be negative. An example of a completely positive description of the robot in this condition is "[it] felt like working together and collaborative."

As for the robot controlled by the human decision-maker, it was described positively using words like "faster/quicker" (21), "responsive/responded" (5), [took] "initiative" (1), "determined" (1), "collaborative" (1), "efficient" (1), "changed behaviour" (1), again used in both completely positive or mixed statements. For example, one participant described the robot with the human decision-maker as "it was faster with short reaction time. It was machine-like and getting the job done." Which is

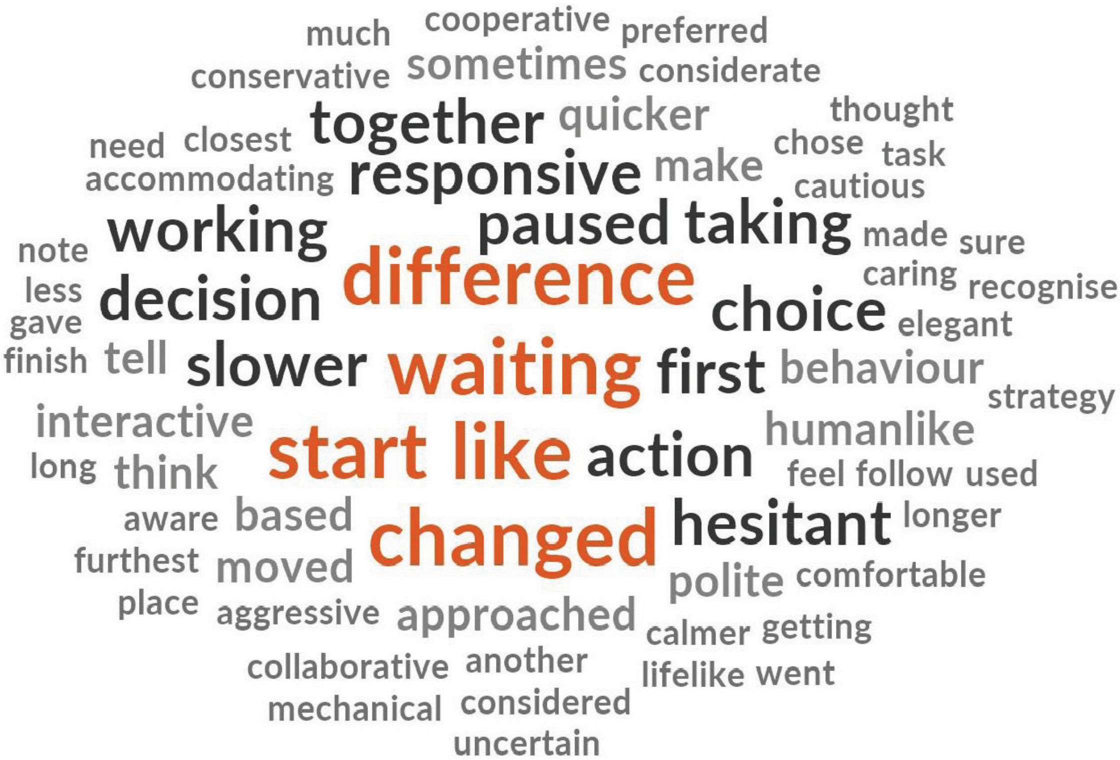

Figure 5.7: NVivo Word Frequency map for the answer to the first interview question for the condition with negotiation layer.

considered a mixed statement as being "machinelike" is taken to be a negative description. An example of a completely positive example is "it was more decisive and quicker."

To show the attitude of the participants when describing each condition, a Matrix Coding Query was performed in NVivo, the result is presented in table 5.14. Attitudes are coded manually in a way that if there are only positive comments about the robot it is coded as "Positive" or if only negative comments are used it is coded as "Negative". If mixed positive and negative adjectives are used it is coded as "Mixed". If there are neither positive nor negative comments about either condition it is coded as "Neutral". In this experiment, the participants' descriptions of the robot were considered neutral when they mention there was no difference between two conditions like saying "couldn't tell any difference", "didn't feel a big difference", or "there was no difference".

Answering the second interview question, 20 participants mentioned they preferred to work with the robot when there was a human decision-maker as it was "quicker". Another 13 participants preferred the condition with the model decision-maker and mentioned the robot was "responsive" (3), "less

Table 5.14: NVivo Matrix Coding Query for Conditions vs Attitudes. Numbers show the number of times participants have described a condition with each attitude.

Matrix Coding Query		Attitudes			
		A : Mixed	B : Negative	C : Neutral	D : Positive
Conditions	Human Decision-Maker	8	3	6	16
	Complete Model	7	7	6	17

aggressive" (1), "quicker" (2), "smoother" (2), "cooperative" (2), "calmer" (1), "more comfortable" (1), and "accommodating" (1) and that's why they preferred it. This is noteworthy that the actual robot movement was the same in both conditions, however, like the first condition, one of the effects of the negotiation layer in the complete model on the robot behaviour was to wait and give way to the participants. 7 participants had no preference between the two conditions. Answering the last question of the interview, 17 participants said the robot with the complete model was more human-like, another 17 participants found the robot with the human decision-maker to be more human-like and 6 of them found no difference between the two conditions. Those choosing the model decision-maker mentioned they found the robot "polite", "kind", "interactive/more interactive", "cooperating" and "responsive" and hence more human-like. While those choosing the human decision-maker described the robot using words like "more efficient", "decisive", "collaborative", "quick", and "taking initiative" and hence more human-like.

5.9 Discussion

It is crystal clear that having a negotiation layer has significantly improved the interaction. All subjective and objective measures were significantly improved in the first Human-Robot Collaboration (HRC) experiment after introducing the negotiation layer (as presented in Section 5.8.1). Although in the qualitative analysis many participants perceived the robot without the negotiation layer to be faster, the actual task completion time for this condition was longer than when using the negotiation layer. It is due to the robot facing many conflicts without the negotiation layer and requiring to repeat its actions leading to a longer task completion time. The absence of the negotiation layer meant there was no hesitation caused due to deliberating on the partner's action and the robot made a decision faster and moved towards a target immediately causing participants to feel it worked faster, while all the robot movements were actually done in the same speed in all conditions. Nonetheless, several participants mentioned the robot with the negotiation layer was more efficient.

The qualitative analysis further revealed that the majority of participants described the robot without the negotiation layer ascribed a negative attitude (Table 5.5) to it, using words like [making] "mistakes", "competitive" and "aggressive", while the robot with the negotiation layer was perceived as "collaborative" and "considerate". Both qualitative and quantitative analyses of the subjective data found that participants consider the robot with the negotiation layer to be more intelligent. The result of the Godspeed questionnaire showed that the robot with the negotiation layer scored better in all categories compared to the robot without the negotiation layer. The score was significantly higher for two categories, namely Anthropomorphism and Perceived Intelligence, further confirming the significant role of the negotiation layer.

Comparing the answers to the second and third post-experiment interview questions in the first HRC experiment, 32 out of 33 participants who preferred the robot with the negotiation layer also considered the robot with the negotiation layer to be more human-like. Only 5 out of 10 participants who preferred the robot without the negotiation layer also considered it more human-like. The other 5 participants still preferred the robot being faster despite not perceiving it as human-like. This indicates the majority of participants preferred to work with a robot with human-like behavioural traits, as they described it is being "considerate", "responsive" or "collaborative".

A further interesting observation in the first HRC experiment is that among 20 participants who experienced the robot without the negotiation layer in their first trial, 12 participants changed the way they worked with the robot from working together in parallel in the first trial to taking turns in the second trial. These participants experienced a high number of conflicts in the first trial with a mean number of conflicts of 6.25 (STD=1.14) which is higher than the overall mean of 5.85 (STD=2.28). This behavioural change was further investigated by looking into the completion time of the first and last characters in the task for both conditions. The character completion time then was divided by the number of blocks per character to have a comparable measure. For the trials without the negotiation layer, the mean time per block goes from 11.435 seconds (STD= 2.94) for the first character to 10.254 seconds (STD= 1.59) for the last character. A one-way ANOVA shows a significant reduction of the time ($F=7.16$, $p<0.0091$). While for the trials with the negotiation layer, the mean time per block goes from 9.538 seconds (STD= 1.482) for the first character to 9.16 seconds (STD= 1.371) for the last character. A one-way ANOVA shows no significant difference between the time for the first and last character ($F=1.37$, $p<0.2446$). This indicates that humans try to adapt and change their behaviour to avoid conflicts in response to adverse or undesirable events in the interaction.

The second HRC experiment was performed to evaluate the decision-making structure compared to a human decision-maker. Overall, the result (as presented in Section 5.8.2) suggests the model performance is close to a human decision-maker considering six participants considered them the same and another 3 participants mentioned they are very similar and there is only a small difference in the speed otherwise they considered them to be the same. Another interesting finding, mentioned in the results, was that an equal number of 17 participants found either condition to be like working with a human. This could be seen in agreement with previous research in autonomous cars by Stanton et al. [110] in which participants were shown video clips of a Tesla Model S P90D when driving on Autopilot and only 50% of participants thought it was controlled by autopilot. Similarly, when driven by a human only 50% of the participants thought it was controlled by a human.

In both HRC experiments, the robot using the complete decision-making model was perceived as slower or calmer than the other conditions namely, without the negotiation layer and human decision-maker. It is due to the robot hesitating when it detects a conflict situation and the negotiation layer deliberates on possible ways to resolve the conflict. This pausing moment did not have an adverse effect on the mean task completion time in the first HRC experiment. In this experiment, despite being perceived as slower, the robot with the negotiation layer has a significantly lower task completion time. This is due to the fact that the robot without the negotiation layer had to repeat its actions after a conflict. However, in the second HRC experiment, the robot with the human decision-maker has a slightly lower task completion time. This difference was not significant and even a further Bayesian comparison analysis of this data showed a high probability of similarity.

In the Second HRC experiment, the human decision-maker performed slightly better than the model in all subjective and objective measures, however, no significant difference was found. To establish similarity further analyses were performed. A Bayesian comparison [12] of the data showed a high probability of similarity for all subjective and objective measures. For the sparse data about the number of conflicts, a comparison of proportion was also performed and this also did not show any difference between the model and the human decision-maker.

Comparing answers to the second and third post-experiment interview questions in the second HRC experiment, 16 out of 20 participants who preferred the robot controlled by a human decision-maker also found the robot more human-like. Similarly, 11 out of 13 participants of those who preferred the robot using the complete model found the robot's behaviour more human-like. Of the 7 participants without any preference for either condition, 2 found the robot with the model more human-like and one participant found the robot with a human decision-maker more human-like. Considering a total of 25 participants preferred the situation where the robot was perceived as being more human-like, it is clear that the approach taken for creating a human-like decision-making model based on human-human interaction could increase the robot's acceptability by the user. However, it is also important to bear in mind that being human-like could be perceived differently by people. Here, the majority of the participants considered a robot, being "calm", "responsive", or "collaborative", more human-like, yet, a small group of participants considered a "faster" robot to be more human-like.

Recorded videos of the three experimental conditions can be seen through the links Without Negotiation, With Negotiation, and Wizard.

CHAPTER 6

MODEL TRANSFERABLILTY

To show the transferability and generic nature of the developed decision-making model, a new experiment is designed. The decision-making models were developed using a generic task. This makes the developed models adaptable to more complex tasks. To demonstrate this ability a new experiment is designed. In this experiment, the task is assembling parts of a toy car. All the developed models could be used in this experiment without requiring retraining. The order of the assembling part was only mapped to the colour order in the colour policy. The experiment is designed similarly to the previous experiments to have two conditions namely, the "Wizard of Oz" and the "Model" for when the robot uses the proposed decision-making architecture. Each participant repeats the task 6 times, 3 for each condition without being informed of the conditions. The order of the trials is randomised in 16 unique combinations (see Table 6.1); hence, 16 participants were recruited.

6.1 Experiment Setup

The experiment setup was similar to the previous human-robot interaction study. Participants were asked to work with a Franka Emika Panda robotic arm in a table-top pick-and-place task. Their dominant hand was marked for tracking by Vicon motion capture reflective balls. The robot control was done in real-time using the libfranka library in C++. The decision-making model was running in parallel in MATLAB and

Table 6.1: Random order of trials for 16 participants. H is for the "Wizard of Oz" condition and M is for the Model condition.

Participant	Trail1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6
1	H	H	H	M	M	M
2	M	M	M	H	H	H
3	H	M	H	M	H	M
4	M	H	M	H	M	H
5	H	H	M	M	H	M
6	M	M	H	H	M	H
7	H	M	M	H	H	M
8	M	H	H	M	M	H
9	M	M	H	H	H	M
10	H	H	M	M	M	H
11	M	H	H	H	M	M
12	H	H	M	H	M	M
13	H	M	M	M	H	H
14	M	M	H	M	H	H
15	H	M	H	H	M	M
16	M	H	M	M	H	H

the decision was communicated to the robot controller through TCP/IP socket communication. This means the robot controller was not continuously receiving the decision but at specific times, hence, the robot could not change its behaviour when one motion was being implemented. Participants repeated the same task in two conditions. The Experiment setup is shown in Figure 6.1. In the human decision-maker condition, the wizard sees the scene from the MS Kinect camera mounted over the table to also track the parts marked by AR markers.

6.2 Wizard Protocol

To guarantee the consistent behaviour of the human decision-maker, the following protocol was given to the wizard to follow: The wizard makes the first decision as soon as the robot is in the "ready position". This is done by entering a number between 1 to 7, 1 for the far-right and 7 for the far-left block from the robot perspective, or entering 8 for just waiting. Later the wizard needs to confirm this decision when the robot goes to the "ready-to-grasp" position. As soon as the robot reaches the "ready-to-grasp" position, the wizard must make the final decision within one second or as soon as the human participant makes his/her move (whichever is faster). At this stage, the wizard has 3 options, either to confirm the

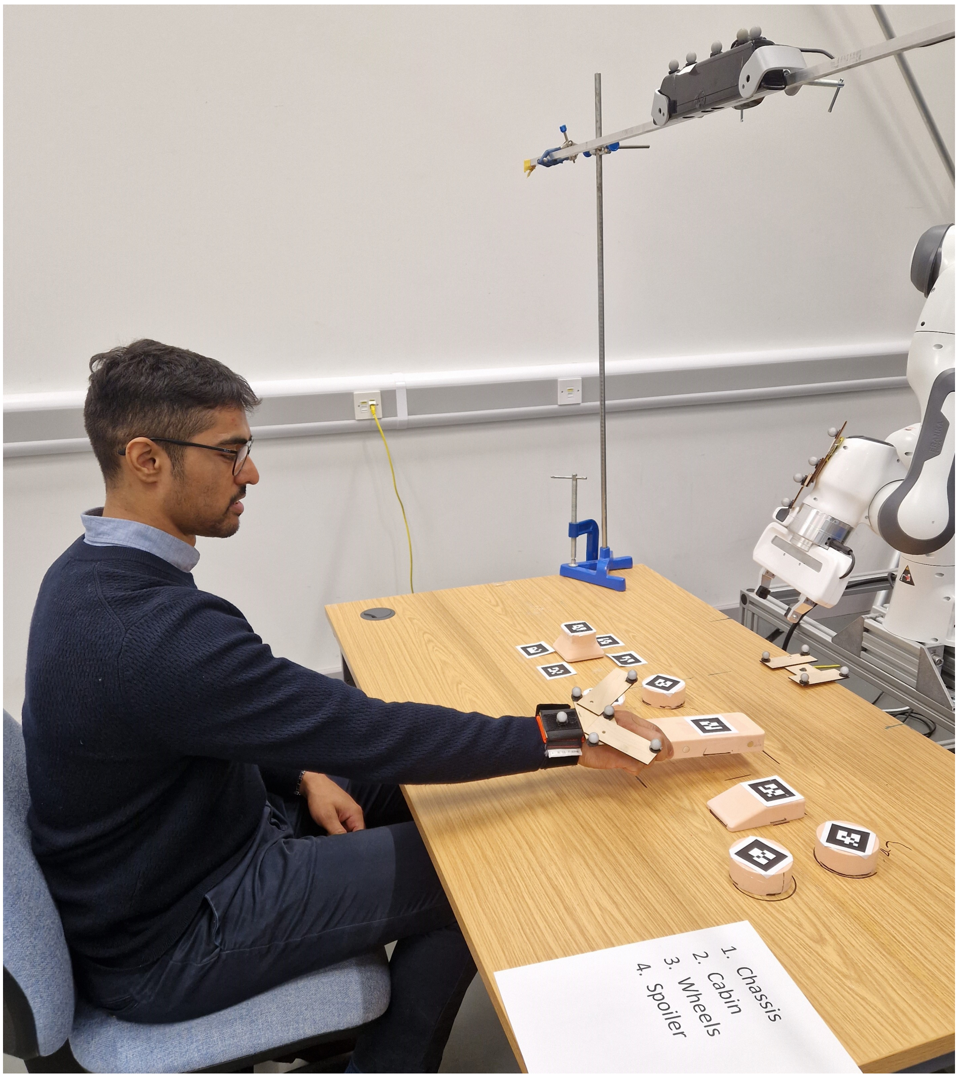

Figure 6.1: Experiment Setup.

decision and continue picking the same block, change to another block, or just wait. If the decision is waiting, the wizard needs to make another decision again within one second until a block is chosen. This will continue until the task is complete. It is noteworthy to mention that the wizard needs to make a decision even if the decision is waiting so that the robot control loop is reset and the robot would be ready for the next action. If no decision is made by the wizard, the robot would stop working.

6.3 Task

The task was to assemble parts of a toy car including the Chassis, Cabin, Wheels and Spoiler. The order of the assembly was given to the participants to follow as 1. Chassis, 2. Cabin, 3. Wheels, and 4. Spoiler. While participants were following this order, they were free to choose any of the wheels without any order when it was the right time to do so. This task was repeated 6 times, 3 times for each condition, in random order. At the end of each trial, participants were asked to answer the PeRDITA questionnaire [35] This questionnaire is chosen over the Godspeed questionnaire in this experiment as it focuses more on the collaboration in the experiment than anthropomorphism of the robot.

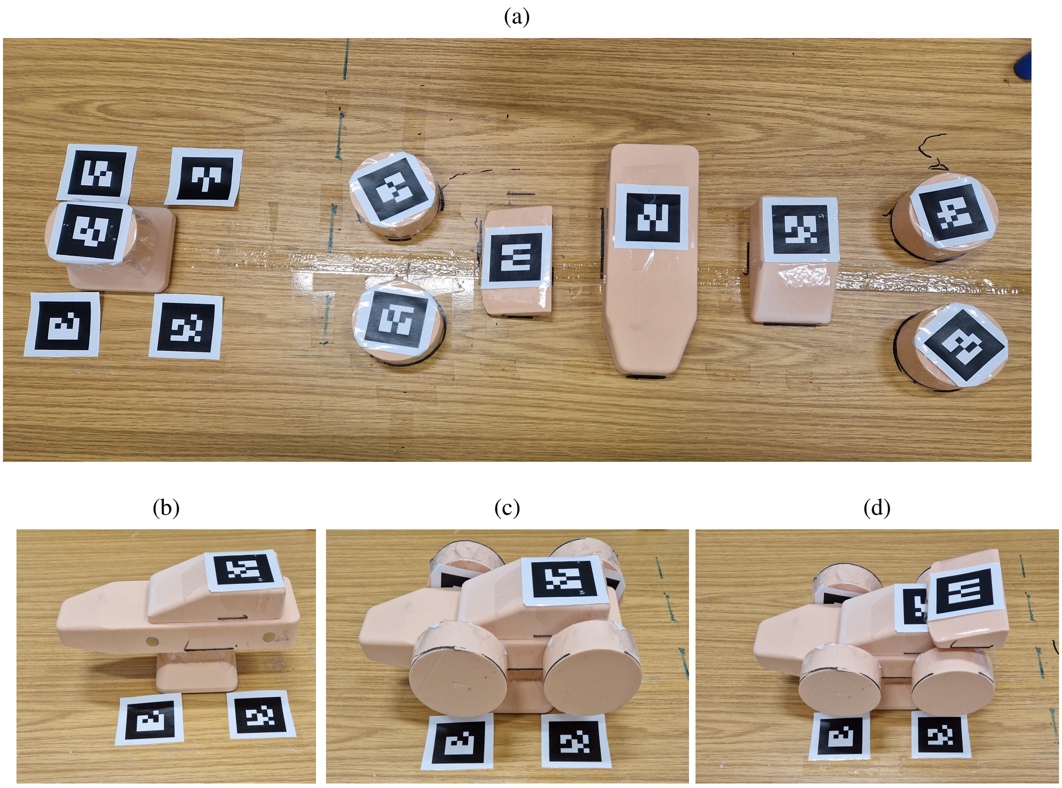

Figure 6.2: Car assembly order: a. Car parts lay over the table on the right and the assembly platform at the left b. First Chassis and then Cabin are assembled c. Wheels assembled next and d. the spoiler is assembled last.

6.4 Participants

In total, 16 participants (13 male) took part in the experiment. Participants were students and staff members of the university with a mean age of 29.8 (STD= 6.14) ranging between 22 and 43 years old with an average height of 175.6 cm (STD=7.43) ranging between 164 and 188 cm. 14 participants were right-handed, one was ambidextrous but used his left hand and one was left-handed, and all reported normal or corrected-to-normal eyesight (8 wearing glasses).

6.5 Objective and Subjective Measures

In terms of objective measures, the task completion time, the robot task share, and the number of conflicts are recorded. The robot task share is calculated as the number of parts the robot picks divided by the total number of parts required (7) for assembling the car. A conflict is considered when the robot tries to pick the part that the participant has just picked or is picking up.

As for the subjective measures, at the end of each trial, participants were asked to answer the PeRDITA questionnaire. As there is no verbal communication between the robot and participants, only four dimensions of the questionnaire were used namely, "Collaboration", "Interaction", "robot Perception", and "Acting".

6.6 Results

The results of the PeRDITA questionnaire are presented in the following tables and graphs. The overall comparison of the results for the human decision-maker against the model is shown in figure 6.3 and Table 6.2 in which the mean is calculated from the average of 3 trials of each condition for each participant for 16 participants. Despite a slightly higher score for the human decision-maker across the board, no significant difference was found in the subjective measures. The results are also presented based on the order of the trials in Figure 6.4 and Table 6.3. Similarly, the data was sorted based on the order of the trials in each condition, namely, Model and Wizard. An overall improvement in scores can be observed, however, no significant difference was found. Bar graph of the mean values for each conditions are presented in Firgue3 6.5 and 6.6.

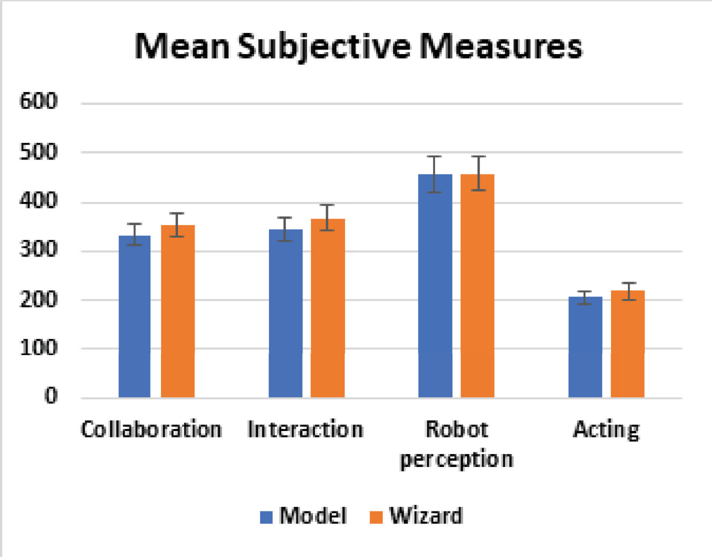

Figure 6.3: Bar graph for Mean value of subjective measures from the PeRDITA questionnaire. Error bars are $\pm 1 SEM$ (Standard Error of Mean)

Table 6.2: Mean and Standard Deviation of Subjective measure from the PeRDITA Questionnaire.

Subjective Measures	Conditions	
	Human Decision-Maker	Model
Collaboration(/500)	333.0208 (STD=91.21448)	352.5 (STD=93.25989)
Interaction(/500)	343.6458 (STD=92.93012)	366.3542 (STD=100.4291)
Robot Perception (/800)	455.4167 (STD=151.6972)	457.8125 (STD=143.0109)
Acting(/300)	206.1458 (STD=51.72119)	217.6042 (STD=66.87736)

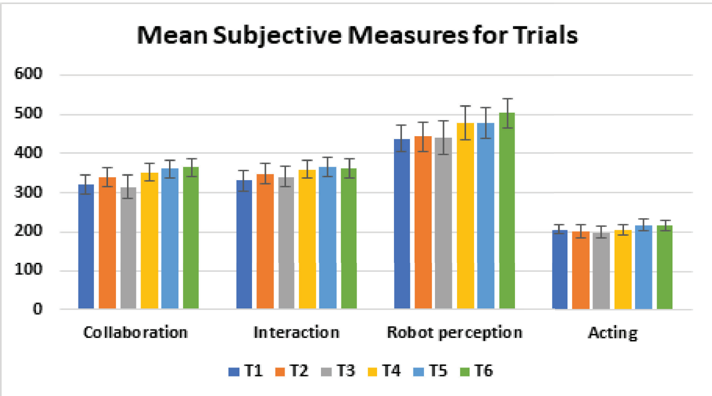

Figure 6.4: Bar graph for Mean value of subjective measures from the PeRDITA questionnaire based on order of the trials. Error bars are $\pm 1 SEM$ (Standard Error of Mean)

Table 6.3: Mean and standard deviation of subjective measures based on the trial order.

Trials	Dimention			
	Collaboration (/500)	Interaction (/500)	Robot perception(/800)	Acting (/300)
T1	321.563 (STD=102.55)	331.875(STD=104.19)	438.75(STD=135.136)	206.875(STD=42.49)
T2	340.625 (STD=101.57)	349.0625 (STD=99.34)	443.4375 (STD=148.8)	200.938 (STD=65.98)
T3	315 (STD=116.56)	340.938 (STD=104.47)	440 (STD= 174.83)	199.063 (STD=59.11)
T4	352.5 (STD=90.4986)	359.375 (STD=90.864)	479.375 (STD=169.14)	205.313 (STD=55.30)
T5	360.625 (STD=90.637)	367.1875 (STD=95.76)	478.438 (STD=161.83)	217.813 (STD=53.35)
T6	365.3125 (STD=90.98)	361.875 (STD=97.96)	502.813 (STD=154.22)	215 (STD=53.45)

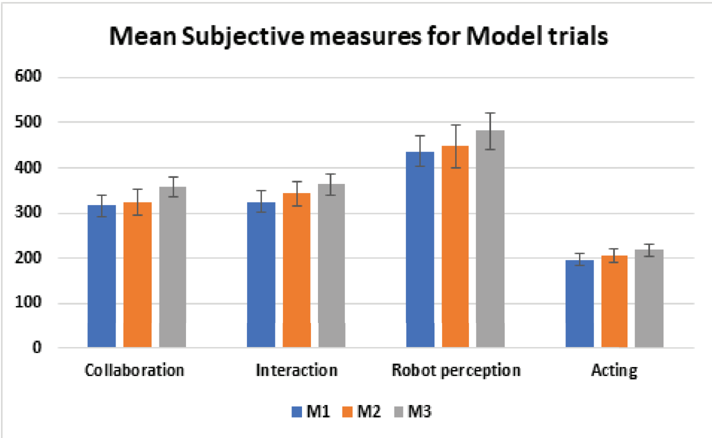

Figure 6.5: Bar graph of mean subjective measures for 3 trials using the model. Error bars are $\pm 1SEM$ (Standard Error of Mean)

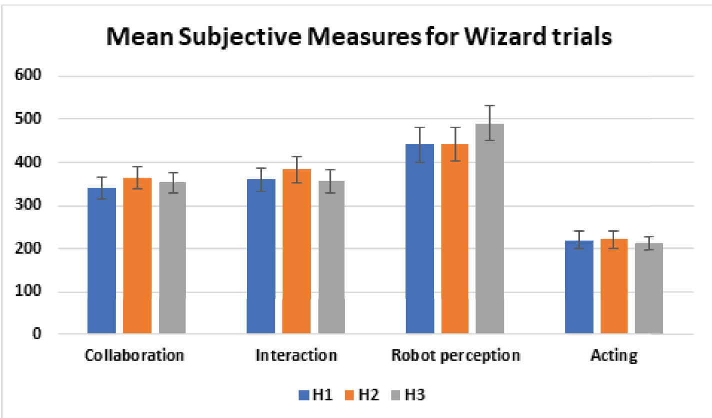

Figure 6.6: Bar graph for Mean Subjective Measures for 3 trials with human decision-maker. Error bars are $\pm 1SEM$ (Standard Error of Mean)

As for objective measures, the data shows a significant difference ($F = 2.57, p\ 0.0323$) in Task completion time between the first and sixth trials (regardless of condition). When data is sorted based on the order of trials within each condition no significant difference was found for the condition using the model but for the human decision-maker, there is a significant difference ($F = 4.04, p$ 0.0243) between the first and third trial in the task completion time. No significant difference was found when comparing the model against the human decision-maker in any of the measures. The results for three objective measures namely, Task Completion Time, Robot Task Share, and Conflicts are presented in Table 6.4 Figures 6.7, 6.8 and 6.9.

Table 6.4: Mean and standard deviation of the objective measure based on conditions.

Assembling Car		Objective Measures		
		Task Completion Time (s)	Robot Task Share	Conflicts
Conditions	Model	54.20833(STD=5.38)	0.44047 (STD=0.0275)	0.145833 (STD=0.171)
	Wizard	53.0833 (STD=5.90856)	0.4583 (STD=0.038391704)	0.08333 (STD=0.1491)

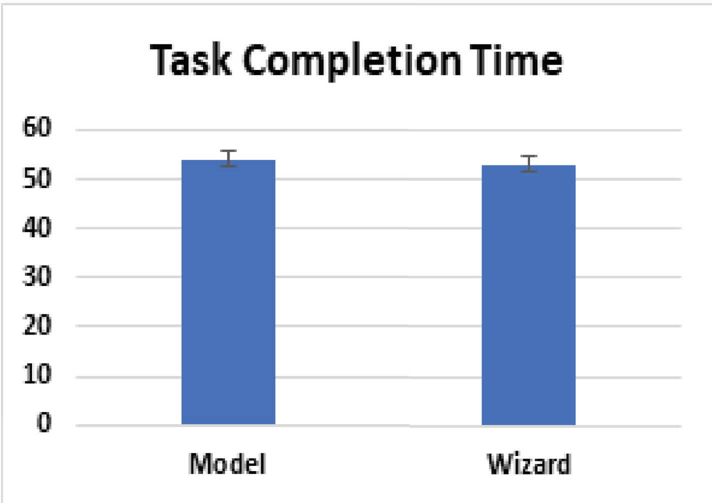

Figure 6.7: Bar graph of mean task completion time based on conditions. Error bars are $\pm 1SEM$ (Standard Error of Mean)

There was a significant difference between the first and last trials in the mean of the Task Completion Time when data was sorted according to the order of the trial both in the overall experiment data and for the Wizard condition. The ANOVA tables are presented in Tables 6.5 and 6.6

The result for the objective measure are also sorted based on the order of the tirals and presented in

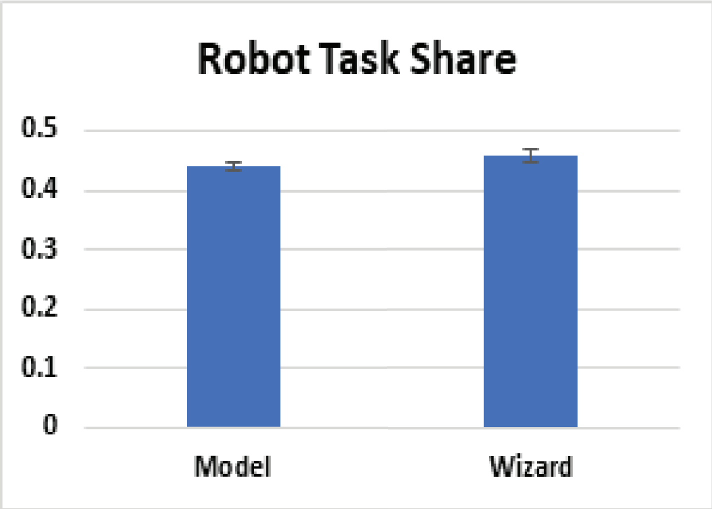

Figure 6.8: Bar graph of mean robot task share based on conditions. Error bars are $\pm 1 SEM$ (Standard Error of Mean)

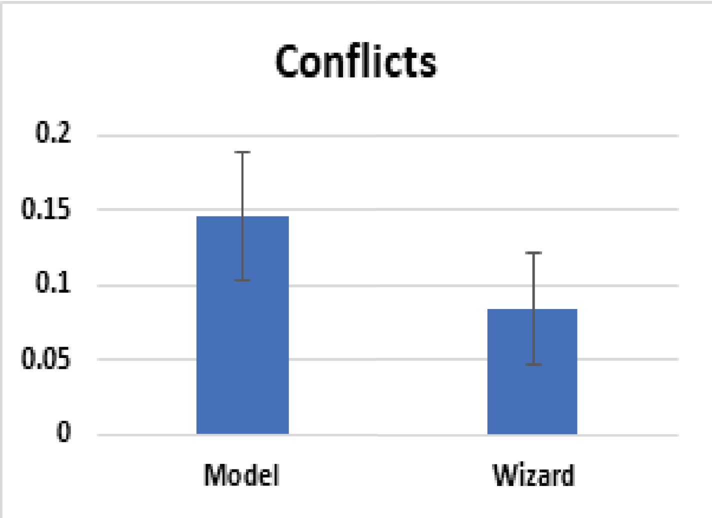

Figure 6.9: Bar graph of mean number of Conflicts based on conditions. Error bars are $\pm 1 SEM$ (Standard Error of Mean)

Table 6.5: ANOVA Table for task completion time based on the order of trials

Task Completion Time					
Source	SS	df	MS	F	Prob > F
Columns	566.46	5	113.292	2.57	0.0323
Error	3973.5	90	44.15		
Total	4339.96	95			

Table 6.7 and Figure 6.10, 6.11 and 6.12.

Table 6.6: ANOVA table for task completion time for human-decision -maker condition based on the order of trials.

Task Completion Time					
Source	SS	df	MS	F	Prob > F
Columns	411.17	2	205.583	4.04	0.0243
Error	2288.5	45	50.856		
Total	2699.67	47			

Table 6.7: Mean and standard deviation of objective measures based on the trial order.

Trials	Objective Measures		
	Task Completion Time	Robot Task Share	Conflicts
T1	56.8125 (STD=8.3364160)	0.446428571 (STD=0.048795)	0.3125 (STD=0.478714)
T2	55.1875 (STD=7.43163284)	0.473214286 (STD=0.06838765)	0.1875 (STD=0.403113)
T3	54.75 (STD=4.892170616)	0.464285714 (STD=0.063888)	0.0625 (STD=0.25)
T4	54.125 (STD=7.098121817)	0.4375 (STD=0.035714286)	0 (STD=0)
T5	51.4375 (STD=6.562202374)	0.4375 (STD=0.035714286)	0.125 (STD=0.341565)
T6	49.5625 (STD=4.774498228)	0.4375 (STD=0.035714286)	0 (STD=0)

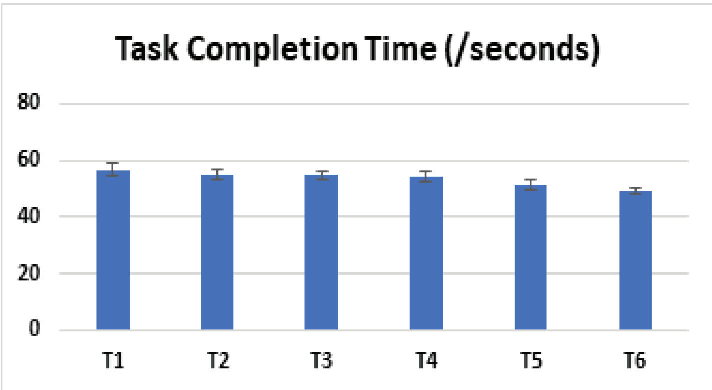

Figure 6.10: Bar graph of mean task completion time (/second) based on the trial order. Error bars are $\pm 1SEM$ (Standard Error of Mean)

Similar to the previous HRC experiment, after finding no significant difference, the Bayesian comparison of the data for two conditions; the "Wizard of Oz" and the "Model", showed a high likelihood of similarity between the two conditions for all subjective and objective measures. The data was normalised for this analysis based on the maximum values of each measure. The comparison results are depicted in Tables 6.8 and 6.9.

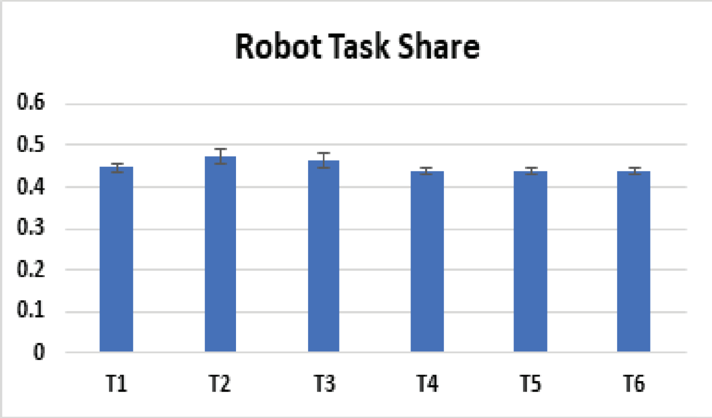

Figure 6.11: Bar graph of mean robot task share based on the trial order. Error bars are $\pm 1SEM$ (Standard Error of Mean)

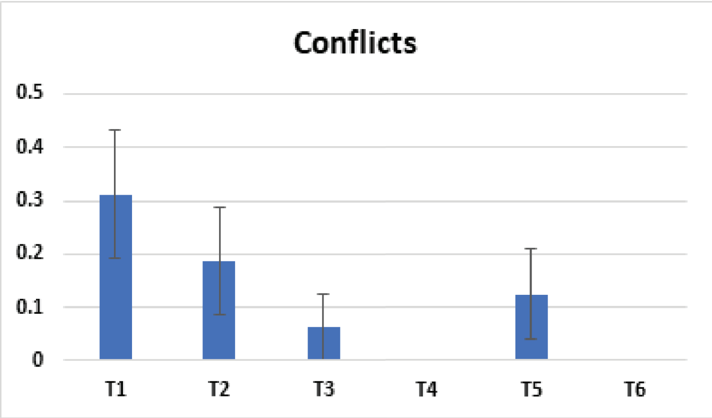

Figure 6.12: Bar graph of mean of the number of the conflicts based on the trial order. Error bars are $\pm 1SEM$ (Standard Error of Mean)

Table 6.8: Bayesian Comparison for the objective measures.

Measure	Model is better	Difference is negligible	Human is better	Rope
Robot Task Share	0.000364	0.9996192	1.680241e-05	0.1
Task Completion Time	1.555164e-05	0.99992125	7.719776e-05	0.1
Conflicts	0.0002701	0.99970587	2.4007e-05	0.5

Table 6.9: Bayesian Comparison for the subjective measures.

Measure	Model is better	Difference is negligible	Human is better	Rope
Collaboration	9.5738e-05	0.975926	0.02397785	0.1
Interaction	8.44871e-05	0.958927	0.040988432	0.1
Robot Perception	0.00182163	0.9954326	0.002745791	0.1
Acting	0.00205497	0.922336	0.0756091	0.1

6.7 Discussion

The car assembling HRC experiment was designed to demonstrate the transferability and expandability of the proposed architecture. To do so, no model was retrained and the order of the assembling task was mapped to already learnt colour order in the colour policy. The distance policy was used as it was trained and indeed proved to be an underlying part of a pick-and-place task.

Bearing in mind that the HRC experiment of assembling the toy car was run with 16 participants, when data was sorted per conditions, the human decision-maker scored slightly better than the model in all the objective and subjective measures but no significant difference was found. A further Bayesian comparison of the data revealed a high probability of similarity between these two conditions.

Using car assembling experiment data sorted on the order of the trials, a significant difference was found in the task completion time between the first and the last trial. The significant difference was also observed when only looking at the human decision-maker trials sorted based on the order of the trials, however, not for the model trials, although the time was improved over trials. The reason for this decrease in the task completion time from the first trial to the last could be that the same task was exactly repeated 6 times during this experiment. This learning effect was prevented in the previous HRC experiments (presented in Chapter 5) as the task was changing randomly for trials and in each trial, participants were making a different alphanumeric character. Having a significant change in the Task Completion Time for the Wizard condition and lack of this significant difference in the model condition could also indicate that participants could adapt better to the robot in the wizard condition. However, further investigation with more participants is required for a concrete conclusion here.

Additionally, fewer conflicts were recorded in the car assembling experiment when compared to the previous HRC experiments. This could be because of having fewer parallel actions in the tasks. As in the previous experiments, in the task used for data collection and training of the models, there were 3 pairs of blocks of the same colours giving them the same priority to be picked in parallel, while in the car assembling task, only wheels have the same priority.

Overall, the result indicates a good potential for transferability of the models trained on a generic task to different more application-specific tasks if they shared some characteristics, such as turn-taking. This is important to note, the Colour policy model was trained for an order of four colours and the car assembling task had also four main steps. However, it is possible to map unequal orders to the "Colour policy" as by only presenting an object the model can create a new stimulus for that object and add it to the order according to the way it was presented to the robot.

A recorded video of the experiment can be seen through this link.

CHAPTER 7

CONCLUSION AND FUTURE WORK

Three research hypotheses were investigated in the work. First, the decision-making process in a joint action for a generic pick-and-place task was modelled. This was done by developing a novel decision-making architecture and by breaking down the decision process into decision policies. Dynamic Neural Field (DNF) was chosen as the modelling framework and it proved to be superior in performance when compared to black-box approaches like Artificial Neural Network (ANN) in our setup. To test the architecture, a generic pick-and-place task was considered in which two decision policies were involved. These policies were chosen carefully to represent both conscious (*Colour policy*) and unconscious (*Distance policy*) cognitive processes involved in the joint action.

The current structure of the decision-making module is designed to be extendable by introducing new policies as needed, which is a clear advantage of the proposed architecture. Currently, it is assumed that both agents have an equal affordance for all actions, however, inspired by the role of canonical neurons, in the future a policy model could be added for the agents' affordances to have an architecture for heterogeneous agents. Another example can be modelling user preferences resulting in a personalised robot that can adapt to the user needs. This makes the architecture suitable for many applications; from production lines to care-working or companion robots for older adults. It is also possible to combine this architecture with the one proposed by Sarthou et al. [101] for considering tasks in which agents have different perspectives of the environment. This could be done either by using their proposed modelling approach or a unified approach of using DNF models for the agent's perspective.

Investigating the second hypothesis, the modelling results indicate that modelling complicated policies can be achieved by integrating single policies and that conflicts can be resolved or prevented in a joint action by means of internal simulation in the proposed Negotiation Layer. The structure can be used for different tasks provided that the relevant policies are modelled and integrated into the system. Partner's action is always considered in the decision-making process for joint actions, hence making this system a good candidate to be embedded in robots for human-robot collaboration (HRC). This was done to further test the architecture and models developed based on the generic task presented in Chapter 3.

Two separate within-subject HRC experiments were conducted to investigate the role of the negotiation layer and compare the architecture to a human decision-maker, respectively. Having the negotiation layer improved all the objective measures significantly as well as Anthropomorphism and Perceived intelligence measures from the Godspeed questionnaire. The "Wizard of Oz" experiments showed that the performance of the proposed decision-making architecture is very close to a human decision-maker as no significant difference was found in any measure and a further Bayesian comparison showed a high probability of similarity.

In both HRC experiments, the robot using the complete decision-making model was perceived as slower or calmer than the other conditions, i.e., without the negotiation layer and human decision-maker. This perceived effect is due to the robot hesitating when it detects a conflict situation and the negotiation layer deliberates on possible ways to resolve the conflict. This pausing moment did not have an adverse effect on the mean task completion time in the first HRC experiment. In this experiment, despite being perceived as slower, the robot with the negotiation layer has a significantly lower task completion time. This is due to the fact that the robot without the negotiation layer had to repeat its actions after a conflict. However, in the second HRC experiment, the robot with the human decision-maker had a slightly lower task completion time. This difference was not significant and even a further Bayesian comparison analysis of this data showed a high probability of similarity. Therefore, overall, the performance of the architecture could be considered similar to a human-decision maker.

One clear shortcoming of the current integration of the architecture with the robot low-level control is that the model is implemented in the MATLAB environment and is communicating with the robot controller through a TCP/IP socket server. This has limited the decision points of the robot so that, if the robot's low-level controller receives a decision from the model and starts implementing it, its action cannot be changed until it is fully implemented.

In an extreme case, this could cause a confusing movement of the robot. For instance, in one case the model outcome was picking one of the last two blocks located on the right side of the scene and while the robot was approaching that block the human partner moved towards the other block at the left of the scene but as the human was moving towards the other block the trajectory passed the block chosen by the robot. This approaching-like movement of the human resulted in the model to change the decision to the other block on the left that human intended to pick but not reached yet. Receiving the changed decision the robot moved towards the other block but by the time it reached the picking position its partner had already picked up the block. This meant the model had to send the robot another change of decision to go back to the previous block on the right. This led to a "zigzag" movement of the robot between two blocks (This was recorded using the ROS tf coordinate frames and a screen-captured video of the ROS visualisation tool "RVIZ" is provided here and the decision change can be seen from second 50 of the video). The lack of real-time communication of the decision to the robot low-level control is also the reason there are conflicts even when the robot is controlled by a human decision-maker. To address this limitation, future work could be to convert the models developed in MATLAB environment to the robot low-level controller environment (c++) so that the model runs continuously in the control loop and can update the robot actions in real-time.

This is noteworthy that exploring individual differences in a joint action was considered out of the scope of this work and hence not explored. However, neuroscientists have investigated the effect of individual personality traits on joint actions. For instance, Novembre et al. [88] found a positive correlation between the individual "perspective-taking" score and their adaptation in the designed joint action task. The "perspective-taking" ability, on the other hand, is known to be stronger in empathic individuals [32, 48, 71]. This led Novembre et al. [88] to use an empathy questionnaire at the end of each trial for each participant. A similar analysis could be done for the presented experiments. Considering few participants, despite facing the robot's "mistakes", preferred to work with the robot that did not take their actions into account and just performed the task, such analysis could reveal any correlation of such preferences with participants' personality traits.

Considering the discrete state of techniques like POMDP, further improvements may be achieved by training the models by means of Partially Observable Monte Carlo Planning (POMCP) as a continuous state method similar to Goldhoorn et al. [52]. Finally, investigating the third research hypothesis, a car assembling task was designed for human-robot collaboration. The car assembling experiment demonstrated the generic nature and expandability of the architecture. The overall architecture performed similarly to the human decision maker in the task using adapted policy models trained for a simple pick-and-place task. Again, *Distance Policy* showed to be an integral part of any task involving the pick-and-place process and its model did not need any retraining. The colour order which was trained for the previous task also was mapped to the order of car assembly and it did not need any retraining. This shows the high potential for the transferability and expandability of the trained models and the architecture for more complex tasks.

In conclusion, three research hypotheses were investigated. Modelling decision policies was done with high accuracy. The presented decision-making architecture demonstrated a human-like performance in HRC experiments with the same task as the one used in the modelling phase and further could be applied to a more realistic task without retraining the models. While the application of the architecture to higher dimension interactions like triadic joint action needs further investigation and is beyond the scope of this work, the results of the modelling and HRC experiments show the potential of the proposed methods for dyadic joint action. Hence, we believe the architecture in its current form could already be applied to many real-world applications like manufacturing.

APPENDIX A

APPENDIX

A.1 Quantitative Graphs and Tables for Two HRC Experiments

Boxplot Graphs and ANOVA tables for the first and second experiments of the objective and subjective measures are presented here.

Table A.1: ANOVA Table for task completion time in the first experiment.

Task Completion Time					
Source	SS	df	MS	F	Prob >F
Columns	61827.2	1	61827.2	17.03	9.15531e-05
Error	283250.8	78	3631.4		
Total	345078	79			

Table A.2: ANOVA Table for robot task share in the first experiment.

Robot Task Share					
Source	SS	df	MS	F	Prob >F
Columns	0.02473	1	0.02473	33.9	1.2273e-07
Error	0.0569	78	0.00073		
Total	0.08164	79			

(a) (b) (c)

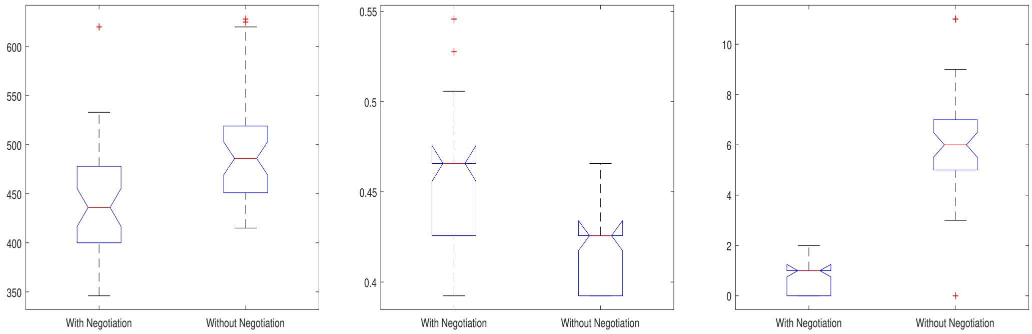

Figure A.1: Boxplots for 3 objective measures: a. Task Completion Time, b. Robot Task Share and c. Number of Conflicts for the first HRC experiment.

Table A.3: ANOVA Table for the number of conflicts in the first HRC experiment.

Conflicts					
Source	SS	df	MS	F	Prob >F
Columns	525.313	1	525.313	205.82	1.39646e-23
Error	199.075	78	2.552		
Total	724.388	79			

(a) (b)

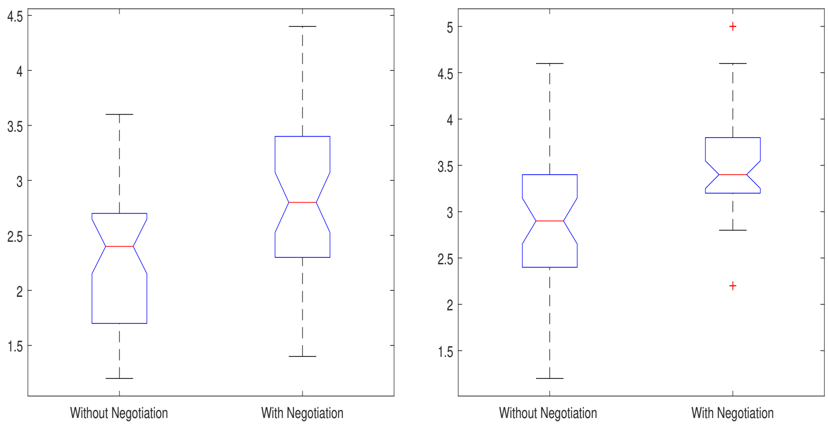

Figure A.2: Boxplots for 2 of subjective measures with significant difference between two conditions: a. Anthropomorphism, b. Perceived Intelligence in the first HRC experiment.

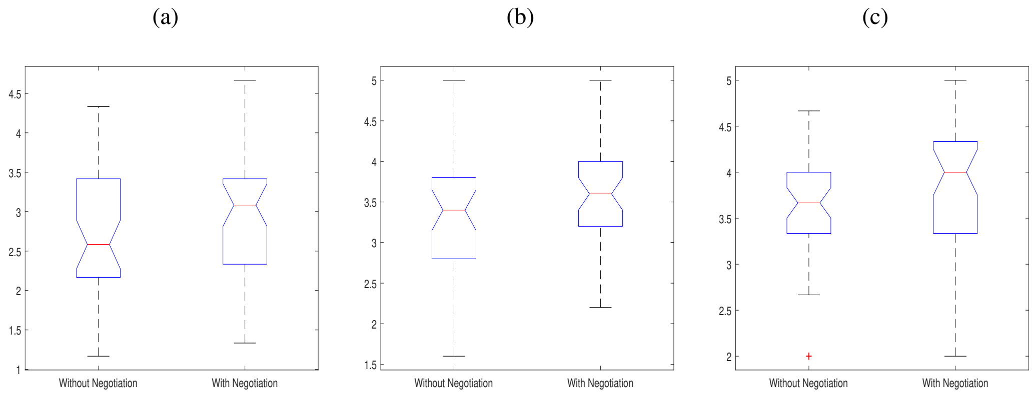

Figure A.3: Boxplots for 3 of subjective measures: a. Animacy, b. Likeability and c. Perceived Safety in the first HRC experiment.

Table A.4: ANOVA Table for the Godspeed Anthropomorphism in the first HRC experiment.

Godspeed Anthropomorphism					
Source	SS	df	MS	F	Prob >F
Columns	5	1	5	8.4	0.0049
Error	46.422	78	0.59515		
Total	51.422	79			

Table A.5: ANOVA Table for the Godspeed Animacy in the first experiment.

Godspeed Animacy					
Source	SS	df	MS	F	Prob > F
Columns	1.7503	1	1.7503	2.93	0.0911
Error	46.6632	78	0.59825		
Total	48.4135	79			

Table A.6: ANOVA Table for the Godspeed Likeability in the first experiment.

Godspeed Likeability					
Source	SS	df	MS	F	Prob > F
Columns	1.682	1	1.682	3.91	0.0516
Error	33.566	78	0.43033		
Total	35.248	79			

Table A.7: ANOVA Table for the Godspeed Perceived Intelligence in the first experiment.

Godspeed Perceived Intelligence					
Source	SS	df	MS	F	Prob > F
Columns	9.248	1	9.248	20.83	1.83802e-o05
Error	34.622	78	0.44387		
Total	43.87	79			

Table A.8: ANOVA Table for the Godspeed Perceived Safety in the first experiment.

Godspeed Perceived Safety					
Source	SS	df	MS	F	Prob > F
Columns	0.6125	1	0.6125	1.51	0.2232
Error	31.6861	78	0.40623		
Total	32.2986	79			

Table A.9: ANOVA Table for task completion time in the second experiment.

Task Completion Time					
Source	SS	df	MS	F	Prob > F
Columns	9990.5	1	9990.45	1.8	0.1841
Error	433846.3	78	5562.13		
Total	443836.8	79			

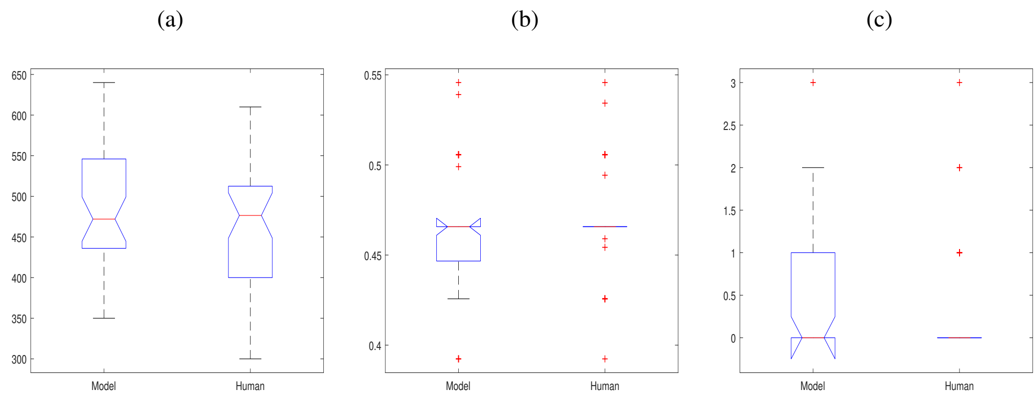

Figure A.4: Boxplots for 3 objective measures: a. Task Completion Time, b. Robot Task Share and c. Number of Conflicts for the second HRC experiment.

Table A.10: ANOVA Table for robot task share in the second experiment.

Robot Task Share					
Source	SS	df	MS	F	Prob > F
Columns	0.00063	1	0.00063	0.56	0.4562
Error	0.08708	78	0.00112		
Total	0.08771	79			

Table A.11: ANOVA Table for the number of conflicts in the second experiment.

Conflicts					
Source	SS	df	MS	F	Prob > F
Columns	1.0125	1	1.0125	1.86	0.1761
Error	42.375	78	0.54327		
Total	43.3875	79			

(a) (b)

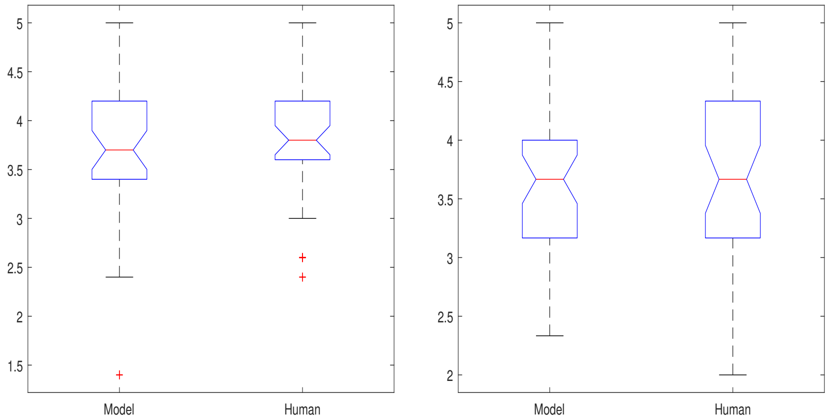

Figure A.5: Boxplots for 2 other subjective measures: a. Perceived Intelligence, b. Perceived Safety in the second HRC experiment.

(a) (b) (c)

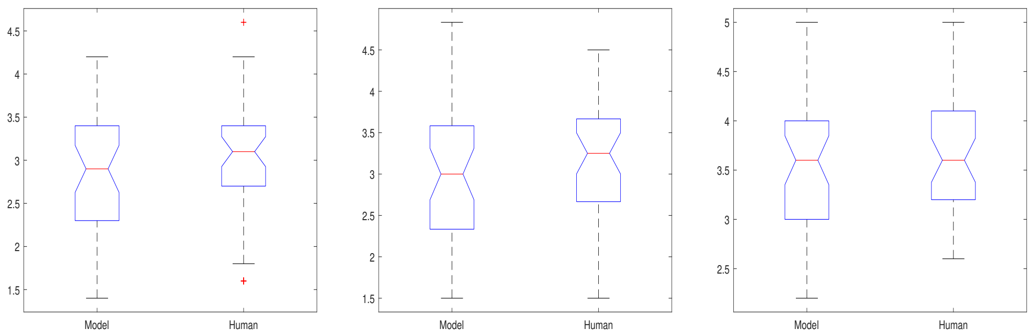

Figure A.6: Boxplots for 3 subjective measures: a. Anthropomorphism, b. Animacy and c. Likeability in the second HRC experiment.

Table A.12: ANOVA Table for the Godspeed Anthropomorphism in the second experiment.

Godspeed Anthropomorphism					
Source	SS	df	MS	F	Prob > F
Columns	0.45	1	0.45	0.95	0.3322
Error	36.87	78	0.47269		
Total	37.32	79			

Table A.13: ANOVA Table for the Godspeed Animacy in the second experiment.

Godspeed Animacy					
Source	SS	df	MS	F	Prob > F
Columns	0.8681	1	0.86806	1.69	0.1978
Error	40.1264	78	0.51444		
Total	40.9944	79			

Table A.14: ANOVA Table for the Godspeed Likeability in the second experiment.

Godspeed Likeability					
Source	SS	df	MS	F	Prob > F
Columns	0.0845	1	0.0845	0.17	0.678
Error	37.935	78	0.48635		
Total	38.0195	79			

Table A.15: ANOVA Table for the Godspeed Perceived Intelligence in the second experiment.

Godspeed Perceived Intelligence					
Source	SS	df	MS	F	Prob > F
Columns	0.2645	1	0.2645	0.6	0.4418
Error	34.515	78	0.4425		
Total	34.7795	79			

Table A.16: ANOVA Table for the Godspeed Perceived Safety in the second experiment.

Godspeed Perceived Safety					
Source	SS	df	MS	F	Prob > F
Columns	0.05	1	0.05	0.09	0.7696
Error	45.1444	78	0.57877		
Total	45.1944	79			

A.2 NVivo Qualitative Tables for Two HRC Experiments

Here the result of the NVivo Word Frequency Query of the answers to the first interview question for two experiments is presented.

Table A.17: NVivo Word Frequency Query of Answers to the first post-experiment interview question describing the robot using the model without negotiation layer in the first experiment.

Word	Length	Count	Weighted Percentage (%)	Similar Words
mistakes	8	12	7.55	mistakes
faster	6	8	5.03	faster
going	5	6	3.77	going
moving	6	5	3.14	moved, moving
aggressive	10	3	1.89	aggressive
quicker	7	3	1.89	quicker
responsive	10	3	1.89	responsive
action	6	2	1.26	action
approach	8	2	1.26	approach, approached
competition	11	2	1.26	competition, competitive
condition	9	2	1.26	condition
confident	9	2	1.26	confident
empty	5	2	1.26	empty
intelligence	12	2	1.26	intelligence, intelligent
times	5	2	1.26	times
wanted	6	2	1.26	wanted
adapt	5	1	0.63	adapt
ahead	5	1	0.63	ahead
alone	5	1	0.63	alone
already	7	1	0.63	already
anticipate	10	1	0.63	anticipate
blocks	6	1	0.63	blocks
collision	9	1	0.63	collision
consideration	13	1	0.63	consideration

Continuation of Table A.17 on the Next Page

Continuation of Table A.17				
decision	8	1	0.63	decision
dominant	8	1	0.63	dominant
dumber	6	1	0.63	dumber
emotional	9	1	0.63	emotional
errors	6	1	0.63	errors
exciting	8	1	0.63	exciting
following	9	1	0.63	following
foolish	7	1	0.63	foolish
getting	7	1	0.63	getting
happen	6	1	0.63	happen
hesitating	10	1	0.63	hesitating
ignored	7	1	0.63	ignored
initiative	10	1	0.63	initiative
intrusive	9	1	0.63	intrusive
irritated	9	1	0.63	irritated
lacked	6	1	0.63	lacked
looking	7	1	0.63	looking
machine	7	1	0.63	machine
making	6	1	0.63	making
movements	9	1	0.63	movements
needs	5	1	0.63	needs
nicer	5	1	0.63	nicer
pleasant	8	1	0.63	pleasant
pressured	9	1	0.63	pressured
puppy	5	1	0.63	puppy
pushing	7	1	0.63	pushing
recognise	9	1	0.63	recognise
repeated	8	1	0.63	repeated
rushing	7	1	0.63	rushing

Continuation of Table A.17 on the Next Page

Continuation of Table A.17				
scared	6	1	0.63	scared
script	6	1	0.63	script
seemed	6	1	0.63	seemed
sensing	7	1	0.63	sensing
several	7	1	0.63	several
smoother	8	1	0.63	smoother
space	5	1	0.63	space
spots	5	1	0.63	spots
sympathetic	11	1	0.63	sympathetic
taking	6	1	0.63	taking
thing	5	1	0.63	thing
trying	6	1	0.63	trying
understand	10	1	0.63	understand
unpredictable	13	1	0.63	unpredictable
useless	7	1	0.63	useless

Table A.18: NVivo Word Frequency Query of Answers to the first post-experiment interview question describing the robot using the complete model with negotiation layer in the first experiment.

Word	Length	Count	Weighted Percentage (%)	Similar Words
waiting	7	11	5.73	waited, waiting
interactive	11	9	4.69	interaction, interactive
slower	6	8	4.17	slower
action	6	6	3.12	action, actions
better	6	5	2.60	better
responsive	10	5	2.60	responsible, responsive
collaborative	13	4	2.08	collaboration, collaborative
aware	5	3	1.56	aware
cared	5	3	1.56	cared, careful, caring
considerate	11	3	1.56	considerate

Continuation of Table A.18 on the Next Page

Continuation of Table A.18				
considering	11	3	1.56	considered, considering
hesitant	8	3	1.56	hesitant
together	8	3	1.56	together
another	7	2	1.04	another
behaviour	9	2	1.04	behaviour
changed	7	2	1.04	changed
condition	9	2	1.04	condition
continuous	10	2	1.04	continuous
easier	6	2	1.04	easier
faster	6	2	1.04	faster
first	5	2	1.04	first
humanlike	9	2	1.04	humanlike
intelligent	11	2	1.04	intelligent
moved	5	2	1.04	moved
predictable	11	2	1.04	predictable
understand	10	2	1.04	understand
accurate	8	1	0.52	accurate
achieve	7	1	0.52	achieve
adaptive	8	1	0.52	adaptive
adjusted	8	1	0.52	adjusted
agile	5	1	0.52	agile
animated	8	1	0.52	animated
apprentice	10	1	0.52	apprentice
backing	7	1	0.52	backing
child	5	1	0.52	child
choices	7	1	0.52	choices
comfortable	11	1	0.52	comfortable
companion	9	1	0.52	companion
compassion	10	1	0.52	compassion

Continuation of Table A.18 on the Next Page

Continuation of Table A.18				
competitive	11	1	0.52	competitive
confident	9	1	0.52	confident
cooperative	11	1	0.52	cooperative
curiosity	9	1	0.52	curiosity
delayed	7	1	0.52	delayed
effective	9	1	0.52	effective
efficient	9	1	0.52	efficient
engaged	7	1	0.52	engaged
enjoyed	7	1	0.52	enjoyed
evenly	6	1	0.52	evenly
finish	6	1	0.52	finish
following	9	1	0.52	following
friendly	8	1	0.52	friendly
going	5	1	0.52	going
human	5	1	0.52	human
impacting	9	1	0.52	impacting
intrigued	9	1	0.52	intrigued
liked	5	1	0.52	liked
longer	6	1	0.52	longer
matched	7	1	0.52	matched
negative	8	1	0.52	negative
opponent	8	1	0.52	opponent
partnership	11	1	0.52	partnership
pattern	7	1	0.52	pattern
perception	10	1	0.52	perception
person	6	1	0.52	person
pleasant	8	1	0.52	pleasant
polite	6	1	0.52	polite
prioritised	11	1	0.52	prioritised

Continuation of Table A.18 on the Next Page

Continuation of Table A.18				
quicker	7	1	0.52	quicker
reading	7	1	0.52	reading
recognised	10	1	0.52	recognised
responding	10	1	0.52	responding
showing	7	1	0.52	showing
situation	9	1	0.52	situation
smooth	6	1	0.52	smooth
smoother	8	1	0.52	smoother
someone	7	1	0.52	someone
something	9	1	0.52	something
steady	6	1	0.52	steady
strategy	8	1	0.52	strategy
synch	5	1	0.52	synch
takes	5	1	0.52	takes
think	5	1	0.52	think
trusted	7	1	0.52	trusted
trying	6	1	0.52	trying
watchful	8	1	0.52	watchful

Table A.19: NVivo Word Frequency Query of Answers to the first post-experiment interview question describing the robot using the complete model with negotiation layer in the second experiment.

Word	Length	Count	Weighted Percentage (%)	Similar Words
waiting	7	21	13.82	waited, waiting, waits
difference	10	7	4.61	difference, different
like	4	6	3.95	like
changed	7	5	3.29	changed
start	5	5	3.29	start
action	6	4	2.63	action
first	5	4	2.63	first

Continuation of Table A.19 on the Next Page

Continuation of Table A.19				
paused	6	4	2.63	pause, paused, pausing
responsive	10	4	2.63	responsive
slower	6	4	2.63	slower
choice	6	3	1.97	choice
decision	8	3	1.97	decision
hesitant	8	3	1.97	hesitant
taking	6	3	1.97	take, taking
together	8	3	1.97	together
working	7	3	1.97	working
approached	10	2	1.32	approached, approaching
based	5	2	1.32	based
behaviour	9	2	1.32	behaviour
humanlike	9	2	1.32	humanlike
interactive	11	2	1.32	interactive
make	4	2	1.32	make, making
moved	5	2	1.32	moved
polite	6	2	1.32	polite
quicker	7	2	1.32	quicker
sometimes	9	2	1.32	sometimes
tell	4	2	1.32	tell
think	5	2	1.32	think, thinking
accommodating	13	1	0.66	accommodating
aggressive	10	1	0.66	aggressive
another	7	1	0.66	another
aware	5	1	0.66	aware
calmer	6	1	0.66	calmer
caring	6	1	0.66	caring
cautious	8	1	0.66	cautious
chose	5	1	0.66	chose

Continuation of Table A.19 on the Next Page

Continuation of Table A.19				
closest	7	1	0.66	closest
collaborative	13	1	0.66	collaborative
comfortable	11	1	0.66	comfortable
conservative	12	1	0.66	conservative
considerate	11	1	0.66	considerate
considered	10	1	0.66	considered
cooperative	11	1	0.66	cooperative
elegant	7	1	0.66	elegant
feel	4	1	0.66	feel
finish	6	1	0.66	finish
follow	6	1	0.66	follow
furthest	8	1	0.66	furthest
gave	4	1	0.66	gave
getting	7	1	0.66	getting
less	4	1	0.66	less
lifelike	8	1	0.66	lifelike
long	4	1	0.66	long
longer	6	1	0.66	longer
made	4	1	0.66	made
mechanical	10	1	0.66	mechanical
much	4	1	0.66	much
need	4	1	0.66	need
note	4	1	0.66	note
place	5	1	0.66	place
preferred	9	1	0.66	preferred
recognise	9	1	0.66	recognise
strategy	8	1	0.66	strategy
sure	4	1	0.66	sure
task	4	1	0.66	task

Continuation of Table A.19 on the Next Page

Continuation of Table A.19				
thought	7	1	0.66	thought
uncertain	9	1	0.66	uncertain
used	4	1	0.66	used
went	4	1	0.66	went

Table A.20: NVivo Word Frequency Query of Answers to the first post-experiment interview question describing the robot with a human decision-maker in the second experiment.

Word	Length	Count	Weighted Percentage (%)	Similar Words
faster	6	15	13.76	faster
difference	10	8	7.34	difference, different
quicker	7	6	5.50	quicker
decision	8	4	3.67	decision, decisive
responsive	10	3	2.75	responsive
waiting	7	3	2.75	waited, waiting
working	7	3	2.75	worked, working
action	6	2	1.83	action
first	5	2	1.83	first
getting	7	2	1.83	getting
interaction	11	2	1.83	interaction, interactive
machinelike	11	2	1.83	machinelike
making	6	2	1.83	making
responded	9	2	1.83	responded, responding
similar	7	2	1.83	similar
based	5	1	0.92	based
behaviour	9	1	0.92	behaviour
changed	7	1	0.92	changed
collaborative	13	1	0.92	collaborative
comparative	11	1	0.92	comparative
competing	9	1	0.92	competing

Continuation of Table A.20 on the Next Page

Continuation of Table A.20				
competitive	11	1	0.92	competitive
complete	8	1	0.92	complete
decide	6	1	0.92	decide
determined	10	1	0.92	determined
efficient	9	1	0.92	efficient
finished	8	1	0.92	finished
follow	6	1	0.92	follow
furthest	8	1	0.92	furthest
independently	13	1	0.92	independently
initiative	10	1	0.92	initiative
liked	5	1	0.92	liked
order	5	1	0.92	order
placing	7	1	0.92	placing
quick	5	1	0.92	quick
reaction	8	1	0.92	reaction
ready	5	1	0.92	ready
rushed	6	1	0.92	rushed
short	5	1	0.92	short
slower	6	1	0.92	slower
smoother	8	1	0.92	smoother
snappier	8	1	0.92	snappier
strategy	8	1	0.92	strategy
taking	6	1	0.92	taking
unspoken	8	1	0.92	unspoken
wanted	6	1	0.92	wanted
wilder	6	1	0.92	wilder

A.3 Questionnaires

GODSPEED I: ANTHROPOMORPHISM

Please rate your impression of the robot on these scales:
以下のスケールに基づいてこのロボットの印象を評価してください。

Fake 偽物のような	1	2	3	4	5	Natural 自然な
Machinelike 機械的	1	2	3	4	5	Humanlike 人間的
Unconscious 意識を持たない	1	2	3	4	5	Conscious 意識を持っている
Artificial 人工的	1	2	3	4	5	Lifelike 生物的
Moving rigidly ぎこちない動き	1	2	3	4	5	Moving elegantly洗練された動き

GODSPEED II: ANIMACY

Please rate your impression of the robot on these scales:
以下のスケールに基づいてこのロボットの印象を評価してください。

Dead 死んでいる	1	2	3	4	5	Alive 生きている
Stagnant 活気のない	1	2	3	4	5	Lively 生き生きとした
Mechanical 機械的な	1	2	3	4	5	Organic 有機的な
Artificial 人工的な	1	2	3	4	5	Lifelike 生物的な
Inert 不活発な	1	2	3	4	5	Interactive 対話的な
Apathetic 無関心な	1	2	3	4	5	Responsive 反応のある

GODSPEED III: LIKEABILITY

Please rate your impression of the robot on these scales:
以下のスケールに基づいてこのロボットの印象を評価してください。

Dislike 嫌い	1	2	3	4	5	Like 好き
Unfriendly 親しみにくい	1	2	3	4	5	Friendly 親しみやすい
Unkind 不親切な	1	2	3	4	5	Kind 親切な
Unpleasant 不愉快な	1	2	3	4	5	Pleasant 愉快な
Awful ひどい	1	2	3	4	5	Nice 良い

GODSPEED IV: PERCEIVED INTELLIGENCE

Please rate your impression of the robot on these scales:
以下のスケールに基づいてこのロボットの印象を評価してください。

Incompetent 無能な	1	2	3	4	5	Competent 有能な
Ignorant 無知な	1	2	3	4	5	Knowledgeable 物知りな
Irresponsible 無責任な	1	2	3	4	5	Responsible 責任のある
Unintelligent 知的でない,	1	2	3	4	5	Intelligent 知的な
Foolish 愚かな	1	2	3	4	5	Sensible 賢明な

GODSPEED V: PERCEIVED SAFETY

Please rate your emotional state on these scales:
以下のスケールに基づいてあなたの心の状態を評価してください。

Anxious 不安な	1	2	3	4	5	Relaxed 落ち着いた
Agitated 動揺している	1	2	3	4	5	Calm 冷静な
Quiescent 平穏な	1	2	3	4	5	Surprised 驚いた

Figure A.7: Godspeed questionnaire [7] used in the HRC experiment presented in chapter 4.

Dimension	Question		
Collaboration	**In your opinion, the collaboration with the robot to perform the task was:**		
	Restrictive		Adaptive
	Useless		Useful
	Unsettling		Satisfactory
	Annoying		Acceptable
	Insecure		Secure
Interaction	**In your opinion, generally, the interaction was:**		
	Negative		Positive
	Complicated		Simple
	Not practical		Practical
	Unpredictable		Predictable
	Ambiguous		Clear
Robot perception	**In your opinion, the robot is rather:**		
	Machinelike		Humanlike
	Artificial		Living
	Inert		Animated
	Apathetic		Responsive
	Unpleasant		Pleasant
	Disagreeable		Agreeable
	Stupid		Intelligent
	Incompetent		Competent
Acting	**In your opinion, the robot actions were:**		
	Inappropriate		Appropriate
	Useless		Useful
	Unpredictable		Predictable

Figure A.8: PeRDITA questionnaire [35] used in the HRC experiment presented in chapter 5.

BIBLIOGRAPHY

[1] D. J. AGRAVANTE, A. CHERUBINI, A. BUSSY, AND A. KHEDDAR, *Human-humanoid joint haptic table carrying task with height stabilization using vision*, in 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2013, pp. 4609–4614.

[2] R. M. ALEXANDER, *A minimum energy cost hypothesis for human arm trajectories*, Biological cybernetics, 76 (1997), pp. 97–105.

[3] S.-I. AMARI, *Dynamics of pattern formation in lateral-inhibition type neural fields*, Biological cybernetics, 27 (1977), pp. 77–87.

[4] J. R. ANDERSON, D. BOTHELL, M. D. BYRNE, S. DOUGLASS, C. LEBIERE, AND Y. QIN, *An integrated theory of the mind.*, Psychological review, 111 (2004), p. 1036.

[5] J. R. ANDERSON, M. MATESSA, AND C. LEBIERE, *Act-r: A theory of higher level cognition and its relation to visual attention*, Human–Computer Interaction, 12 (1997), pp. 439–462.

[6] W. A. BAINBRIDGE, J. HART, E. S. KIM, AND B. SCASSELLATI, *The effect of presence on human-robot interaction*, in RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication, 2008, pp. 701–706.

[7] C. BARTNECK, D. KULIĆ, E. CROFT, AND S. ZOGHBI, *Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots*, International journal of social robotics, 1 (2009), pp. 71–81.

[8] A. BAUER, D. WOLLHERR, AND M. BUSS, *Human–robot collaboration: a survey*, International Journal of Humanoid Robotics, 5 (2008), pp. 47–66.

[9] G. M. BECKER, M. H. DEGROOT, AND J. MARSCHAK, *Probabilities of choices among very similar objects: An experiment to decide between two models*, Behavioral Science, 8 (1963), pp. 306–311.

[10] H. Bekkering, E. R. De Bruijn, R. H. Cuijpers, R. Newman-Norlund, H. T. Van Schie, and R. Meulenbroek, *Joint action: Neurocognitive mechanisms supporting human interaction*, Topics in Cognitive Science, 1 (2009), pp. 340–352.

[11] R. Bellman, *A markovian decision process*, tech. rep., DTIC Document, 1957.

[12] A. Benavoli, G. Corani, J. Demšar, and M. Zaffalon, *Time for a change: a tutorial for comparing multiple classifiers through bayesian analysis*, Journal of Machine Learning Research, 18 (2017), pp. 1–36.

[13] G. Beraldo, L. Tonin, J. d. R. Millán, and E. Menegatti, *Shared intelligence for robot teleoperation via bmi*, IEEE Transactions on Human-Machine Systems, 52 (2022), pp. 400–409.

[14] L. Bi, A. G. Feleke, and C. Guan, *A review on emg-based motor intention prediction of continuous human upper limb motion for human-robot collaboration*, Biomedical Signal Processing and Control, 51 (2019), pp. 113–127.

[15] F. Bianco and D. Ognibene, *Functional advantages of an adaptive theory of mind for robotics: a review of current architectures*, in 2019 11th Computer Science and Electronic Engineering (CEEC), 2019, pp. 139–143.

[16] E. Bicho, W. Erlhagen, L. Louro, and E. C. e Silva, *Neuro-cognitive mechanisms of decision making in joint action: A human–robot interaction study*, Human movement science, 30 (2011), pp. 846–868.

[17] E. Bicho, L. Louro, and W. Erlhagen, *Integrating verbal and nonverbal communication in a dynamic neural field architecture for human-robot interaction*, Frontiers in neurorobotics, 4 (2010), p. 5.

[18] M. J. Blanca Mena, R. Alarcón Postigo, J. Arnau Gras, R. Bono Cabré, R. Bendayan, et al., *Non-normal data: Is anova still a valid option?*, Psicothema, (2017).

[19] C. Brosque, E. Galbally, O. Khatib, and M. Fischer, *Human-robot collaboration in construction: Opportunities and challenges*, in 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2020, pp. 1–8.

[20] G. Buisan, A. Favier, A. Mayima, and R. Alami, *Hatp/ehda: A robot task planner anticipating and eliciting human decisions and actions*, in 2022 International Conference on Robotics and Automation (ICRA), 2022, pp. 2818–2824.

[21] J. R. Busemeyer and J. T. Townsend, *Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment.*, Psychological review, 100 (1993), p. 432.

[22] A. R. Cassandra, *A survey of pomdp applications*, in Working Notes of AAAI 1998 Fall Symposium on Planning with Partially Observable Markov Decision Processes, vol. 1724, Citeseer, 1998.

[23] I. Chadès, L. V. Pascal, S. Nicol, C. S. Fletcher, and J. Ferrer-Mestres, *A primer on partially observable markov decision processes (pomdps)*, Methods in Ecology and Evolution, 12 (2021), pp. 2058–2072.

[24] K. Chandan, V. Kudalkar, X. Li, and S. Zhang, *Negotiation-based human-robot collaboration via augmented reality*, arXiv preprint arXiv:1909.11227, (2019).

[25] C. Chao and A. L. Thomaz, *Timing in multimodal turn-taking interactions: Control and analysis using timed petri nets*, Journal of Human-Robot Interaction, 1 (2011), pp. 1–16.

[26] Y. Cheng, P. Yi, R. Liu, J. Dong, D. Zhou, and Q. Zhang, *Human-robot interaction method combining human pose estimation and motion intention recognition*, in 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2021, pp. 958–963.

[27] H. Choset, K. M. Lynch, S. Hutchinson, G. A. Kantor, and W. Burgard, *Principles of robot motion: theory, algorithms, and implementations*, MIT press, 2005.

[28] N. Churamani, P. Barros, H. Gunes, and S. Wermter, *Affect-driven learning of robot behaviour for collaborative human-robot interactions*, Frontiers in Robotics and AI, (2022), p. 20.

[29] Cosiniva, *Cosiniva: Dynamic field theory matlab toolbox.* https://dynamicfieldtheory.org/cosivina/, 2019-02-05.

[30] D. R. Cox and H. D. Miller, *The theory of stochastic processes*, vol. 134, CRC Press, 1977.

[31] A. Curioni, G. Knoblich, and N. Sebanz, *Joint action in humans: A model for human-robot interactions*, in Humanoid Robotics: A Reference, A. Goswami and P. Vadakkepat, eds., Springer Nature, Switzerland, 2016, pp. 1–19.

[32] J. Decety and H. Jackson, *The functional architecture of human empathy. behavioural and cognitive neuroscience review, 3 (2), 71-100*, 2004.

[33] S. DEVIN AND R. ALAMI, *An implemented theory of mind to improve human-robot shared plans execution*, in 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2016, pp. 319–326.

[34] ——, *An implemented theory of mind to improve human-robot shared plans execution*, in 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), IEEE, 2016, pp. 319–326.

[35] S. DEVIN, C. VRIGNAUD, K. BELHASSEIN, A. CLODIC, O. CARRERAS, AND R. ALAMI, *Evaluating the pertinence of robot decisions in a human-robot joint action context: The perdita questionnaire*, in 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2018, pp. 144–151.

[36] U. DOMBROWSKI, T. STEFANAK, AND J. PERRET, *Interactive simulation of human-robot collaboration using a force feedback device*, Procedia Manufacturing, 11 (2017), pp. 124–131.

27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy.

[37] V. DUTTA AND T. ZIELINSKA, *Predicting the intention of human activities for real-time human-robot interaction (hri)*, in Social Robotics, A. Agah, J.-J. Cabibihan, A. M. Howard, M. A. Salichs, and H. He, eds., Cham, 2016, Springer International Publishing, pp. 723–734.

[38] W. ERLHAGEN, *Internal models for visual perception*, Biological cybernetics, 88 (2003), pp. 409–417.

[39] W. ERLHAGEN AND E. BICHO, *A dynamic neural field approach to natural and efficient human-robot collaboration*, (2014), pp. 341–365.

[40] W. ERLHAGEN, A. MUKOVSKIY, AND E. BICHO, *A dynamic model for action understanding and goal-directed imitation*, Brain research, 1083 (2006), pp. 174–188.

[41] W. ERLHAGEN, A. MUKOVSKIY, E. BICHO, G. PANIN, C. KISS, A. KNOLL, H. VAN SCHIE, AND H. BEKKERING, *Goal-directed imitation for robots: A bio-inspired approach to action understanding and skill learning*, Robotics and autonomous systems, 54 (2006), pp. 353–360.

[42] X. FAN, S. SUN, AND J. YEN, *On shared situation awareness for supporting human decision-making teams.*, in AAAI Spring Symposium: AI Technologies for Homeland Security, 2005, pp. 17–24.

[43] X. FAN AND J. YEN, *R-cast: Integrating team intelligence for human-centered teamwork*, in Proceedings of the national conference on artificial intelligence, vol. 22, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2007, p. 1535.

[44] C. FAUBEL AND G. SCHÖNER, *Learning to recognize objects on the fly: a neurally based dynamic field approach*, Neural networks, 21 (2008), pp. 562–576.

[45] A. FOKA AND P. TRAHANIAS, *Real-time hierarchical pomdps for autonomous robot navigation*, Robotics and Autonomous Systems, 55 (2007), pp. 561–571.

[46] S. FUCHS AND A. BELARDINELLI, *Gaze-based intention estimation for shared autonomy in pick-and-place tasks*, Frontiers in Neurorobotics, 15 (2021), p. 647930.

[47] V. GALLESE AND A. GOLDMAN, *Mirror neurons and the simulation theory of mind-reading*, Trends in cognitive sciences, 2 (1998), pp. 493–501.

[48] V. GAZZOLA, L. AZIZ-ZADEH, AND C. KEYSERS, *Empathy and the somatotopic auditory mirror system in humans*, Current biology, 16 (2006), pp. 1824–1829.

[49] M. A. GIESE, *Dynamic neural field theory for motion perception*, vol. 469, Springer Science & Business Media, New York, 2012.

[50] N. GILDERT, A. G. MILLARD, A. POMFRET, AND J. TIMMIS, *The need for combining implicit and explicit communication in cooperative robotic systems*, Frontiers in Robotics and AI, 5 (2018), p. 65.

[51] G. V. GLASS, P. D. PECKHAM, AND J. R. SANDERS, *Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance*, Review of educational research, 42 (1972), pp. 237–288.

[52] A. GOLDHOORN, A. GARRELL, R. ALQUÉZAR, AND A. SANFELIU, *Searching and tracking people with cooperative mobile robots*, Autonomous Robots, 42 (2018), pp. 739–759.

[53] R. M. GORDON, *Folk psychology as simulation*, Mind & language, 1 (1986), pp. 158–171.

[54] O. C. GÖRÜR, B. ROSMAN, F. SIVRIKAYA, AND S. ALBAYRAK, *Social cobots: Anticipatory decision-making for collaborative robots incorporating unexpected human behaviors*, in Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, 2018, pp. 398–406.

[55] O. C. GÖRÜR, B. S. ROSMAN, G. HOFFMAN, AND S. ALBAYRAK, *Toward integrating theory of mind into adaptive decision-making of social robots to understand human intention*, (2017).
accessed at: https://researchspace.csir.co.za/dspace/bitstream/handle/10204/9653/G

[56] R. GROTEN, D. FETH, R. L. KLATZKY, AND A. PEER, *The role of haptic feedback for the integration of intentions in shared task execution*, IEEE Transactions on Haptics, 6 (2013), pp. 94–105.

[57] S. GRUSHKO, A. VYSOCKỲ, D. HECZKO, AND Z. BOBOVSKỲ, *Intuitive spatial tactile feedback for better awareness about robot trajectory during human–robot collaboration*, Sensors, 21 (2021), p. 5748.

[58] J. GUITTON, M. WARNIER, AND R. ALAMI, *Belief management for hri planning*, in European Conference on Artificial Intelligence-Workshop on Belief change, Non-monotonic reasoning and Conflict Resolution BNC@ ECAI 2012, 2012.

[59] Y. GUO, Y. LIU, A. OERLEMANS, S. LAO, S. WU, AND M. S. LEW, *Deep learning for visual understanding: A review*, Neurocomputing, 187 (2016), pp. 27–48.

[60] Y. Y. HAIMES, D. LI, AND V. TULSIANI, *Multiobjective decision-tree analysis 1*, Risk Analysis, 10 (1990), pp. 111–127.

[61] M. R. HARWELL, E. N. RUBINSTEIN, W. S. HAYES, AND C. C. OLDS, *Summarizing monte carlo results in methodological research: The one-and two-factor fixed effects anova cases*, Journal of educational statistics, 17 (1992), pp. 315–339.

[62] D. M. HAUSMAN, *The handbook of economic methodology, john davis, d. wade hands, and uskali mäki (eds.). edward elgar, 1998, xviii+ 572 pages.*, Economics and Philosophy, 15 (1999), pp. 289–295.

[63] H.-P. HUANG AND C.-C. LIANG, *Strategy-based decision making of a soccer robot system using a real-time self-organizing fuzzy decision tree*, Fuzzy Sets and Systems, 127 (2002), pp. 49–64.

[64] M. IACOBONI, I. MOLNAR-SZAKACS, V. GALLESE, G. BUCCINO, J. C. MAZZIOTTA, AND G. RIZZOLATTI, *Grasping the intentions of others with one's own mirror neuron system*, PLoS biology, 3 (2005), p. e79.

[65] C. IGEL, W. ERLHAGEN, AND D. JANCKE, *Optimization of dynamic neural fields*, Neurocomputing, 36 (2001), pp. 225–233.

[66] S. JAIN AND B. ARGALL, *Probabilistic human intent recognition for shared autonomy in assistive robotics*, ACM Transactions on Human-Robot Interaction (THRI), 9 (2019), pp. 1–23.

[67] D. JANCKE, W. ERLHAGEN, H. R. DINSE, A. C. AKHAVAN, M. GIESE, A. STEINHAGE, AND G. SCHÖNER, *Parametric population representation of retinal location: Neuronal interaction dynamics in cat primary visual cortex*, The Journal of Neuroscience, 19 (1999), pp. 9016–9028.

[68] Y. JI, F. XU, Y. YANG, F. SHEN, H. T. SHEN, AND W.-S. ZHENG, *A large-scale rgb-d database for arbitrary-view human action recognition*, in Proceedings of the 26th ACM international Conference on Multimedia, 2018, pp. 1510–1518.

[69] L. JOHANNSMEIER AND S. HADDADIN, *A hierarchical human-robot interaction-planning framework for task allocation in collaborative industrial assembly processes*, IEEE Robotics and Automation Letters, 2 (2016), pp. 41–48.

[70] D. KAHNEMAN AND A. TVERSKY, *Prospect theory: An analysis of decision under risk*, Econometrica: Journal of the econometric society, (1979), pp. 263–291.

[71] J. T. KAPLAN AND M. IACOBONI, *Getting a grip on other minds: Mirror neurons, intention understanding, and cognitive empathy*, Social neuroscience, 1 (2006), pp. 175–183.

[72] J. E. LAIRD, *Extending the soar cognitive architecture*, Frontiers in Artificial Intelligence and Applications, 171 (2008), p. 224.

[73] ———, *The Soar cognitive architecture*, MIT press, Cambridge, Massachusetts, 2012.

[74] ———, *An analysis and comparison of act-r and soar*, arXiv preprint arXiv:2201.09305, (2022).

[75] J. E. LAIRD, K. R. KINKADE, S. MOHAN, AND J. Z. XU, *Cognitive robotics using the soar cognitive architecture.*, in CogRob@ AAAI, Citeseer, 2012.

[76] P. LANGLEY, J. E. LAIRD, AND S. ROGERS, *Cognitive architectures: Research issues and challenges*, Cognitive Systems Research, 10 (2009), pp. 141–160.

[77] K. M. LEE, Y. JUNG, J. KIM, AND S. R. KIM, *Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction*, International Journal of Human-Computer Studies, 64 (2006), pp. 962–973.

[78] L. M. LIX, J. C. KESELMAN, AND H. J. KESELMAN, *Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance f test*, Review of educational research, 66 (1996), pp. 579–619.

[79] M. LOHSE, *The role of expectations and situations in human-robot interaction*, New Frontiers in Human-Robot Interaction, (2011), pp. 35–56.

[80] G. METTA, G. SANDINI, L. NATALE, L. CRAIGHERO, AND L. FADIGA, *Understanding mirror neurons: a bio-robotic approach*, Interaction studies, 7 (2006), pp. 197–232.

[81] M. MITCHELL, *An introduction to genetic algorithms*, MIT press, Cambridge, Massachusetts, 1998.

[82] R. MORATZ AND T. TENBRINK, *Affordance-based human-robot interaction*, in Towards Affordance-Based Robot Control, Springer, 2008, pp. 63–76.

[83] R. R. MURPHY, *Introduction to AI robotics*, MIT press, 2019.

[84] B. MUTLU, A. TERRELL, AND C.-M. HUANG, *Coordination mechanisms in human-robot collaboration*, in Proceedings of the Workshop on Collaborative Manipulation, 8th ACM/IEEE International Conference on Human-Robot Interaction, Citeseer, 2013, pp. 1–6.

[85] A. MÖRTL, M. LAWITZKY, A. KUCUKYILMAZ, M. SEZGIN, C. BASDOGAN, AND S. HIRCHE, *The role of roles: Physical cooperation between humans and robots*, The International Journal of Robotics Research, 31 (2012), pp. 1656–1674.

[86] E. NG, Z. LIU, AND M. KENNEDY, *It takes two: Learning to plan for human-robot cooperative carrying*, in 2023 IEEE International Conference on Robotics and Automation (ICRA), 2023, pp. 7526–7532.

[87] R. M. NOSOFSKY AND T. J. PALMERI, *An exemplar-based random walk model of speeded classification.*, Psychological review, 104 (1997), p. 266.

[88] G. NOVEMBRE, L. F. TICINI, S. SCHÜTZ-BOSBACH, AND P. E. KELLER, *Motor simulation and the coordination of self and other in real-time joint action*, Social cognitive and affective neuroscience, 9 (2014), pp. 1062–1068.

[89] R. OLIVAS, *Decision trees*, A Primer for Decision-making Professionals, (2007).

[90] E. PACHERIE AND J. DOKIC, *From mirror neurons to joint actions*, Cognitive Systems Research, 7 (2006), pp. 101–112.

[91] G. I. PARISI, *Human Action Recognition and Assessment Via Deep Neural Network Self-Organization*, Springer International Publishing, Cham, 2020, pp. 187–211.

[92] J. PINEAU AND G. J. GORDON, *Pomdp planning for robust robot control*, (2007), pp. 69–82.

[93] S. C. O. S. W. PNG AND D. H. W. S. LEE, *Pomdps for robotic tasks with mixed observability*, (2009).

[94] R. POPPE, *A survey on vision-based human action recognition*, Image and vision computing, 28 (2010), pp. 976–990.

[95] I. RISTEA, *Reflections on mechanisms influencing human behavior*, Procedia-Social and Behavioral Sciences, 92 (2013), pp. 799–805.

[96] F. E. RITTER, F. TEHRANCHI, AND J. D. OURY, *Act-r: A cognitive architecture for modeling cognition*, Wiley Interdisciplinary Reviews: Cognitive Science, 10 (2019), p. e1488.

[97] R. M. ROE, J. R. BUSEMEYER, AND J. T. TOWNSEND, *Multialternative decision field theory: A dynamic connectionst model of decision making.*, Psychological review, 108 (2001), pp. 370–392.

[98] J. K. ROSENBLATT, *Optimal selection of uncertain actions by maximizing expected utility*, Autonomous Robots, 9 (2000), pp. 17–25.

[99] Y. SANDAMIRSKAYA, *Dynamic neural fields as a step toward cognitive neuromorphic architectures*, Frontiers in neuroscience, 7 (2014), p. 276.

[100] Y. SANDAMIRSKAYA AND G. SCHÖNER, *An embodied account of serial order: How instabilities drive sequence generation*, Neural Networks, 23 (2010), pp. 1164–1179.

[101] G. SARTHOU, A. MAYIMA, G. BUISAN, K. BELHASSEIN, AND A. CLODIC, *The director task: a psychology-inspired task to assess cognitive and interactive robot architectures*, in 2021 30th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2021, pp. 770–777.

[102] G. SCHÖNER, *Dynamical systems approaches to cognition*, in The Cambridge Handbook of Computational Psychology, R. Sun, ed., Cambridge University Press, 2008, pp. 101–126.

[103] G. SCHÖNER, M. DOSE, AND C. ENGELS, *Dynamics of behavior: Theory and applications for autonomous robot architectures*, Robotics and autonomous systems, 16 (1995), pp. 213–245.

[104] N. SEBANZ AND G. KNOBLICH, *Prediction in joint action: What, when, and where*, Topics in Cognitive Science, 1 (2009), pp. 353–367.

[105] K. SHANTON AND A. GOLDMAN, *Simulation theory*, Wiley Interdisciplinary Reviews: Cognitive Science, 1 (2010), pp. 527–538.

[106] W. SHENG, A. THOBBI, AND Y. GU, *An integrated framework for human–robot collaborative manipulation*, IEEE transactions on cybernetics, 45 (2014), pp. 2030–2041.

[107] P. L. SMITH, *Psychophysically principled models of visual simple reaction time.*, Psychological review, 102 (1995), p. 567.

[108] J. A. SOKOLOWSKI, *Can a composite agent be used to implement a recognition-primed decision model*, in Proceedings of the Eleventh Conference on Computer Generated Forces and Beharioral Representation, 2002, pp. 431–436.

[109] M. T. SPAAN AND N. SPAAN, *A point-based pomdp algorithm for robot planning*, in Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on, vol. 3, IEEE, 2004, pp. 2399–2404.

[110] N. A. STANTON, A. ERIKSSON, V. A. BANKS, AND P. A. HANCOCK, *Turing in the driver's seat: Can people distinguish between automated and manually driven vehicles?*, Human Factors and Ergonomics in Manufacturing & Service Industries, 30 (2020), pp. 418–425.

[111] E. SWERE AND D. J. MULVANEY, *Robot navigation using decision trees*, Electronic systems and control division research, (2003).

[112] K. A. TAHBOUB, *Intelligent human-machine interaction based on dynamic bayesian networks probabilistic intention recognition*, Journal of Intelligent and Robotic Systems, 45 (2006), pp. 31–52.

[113] Y. TANIAI AND J. NISHII, *Optimality of reaching movements based on energetic cost under the influence of signal-dependent noise*, in Neural Information Processing: 14th International Conference, ICONIP 2007, Kitakyushu, Japan, November 13-16, 2007, Revised Selected Papers, Part I 14, Springer, 2008, pp. 1091–1099.

[114] J. T. TOWNSEND AND J. R. BUSEMEYER, *Approach-avoidance: Return to dynamic decision behavior*, in Current Issues in Cognitive Processes: The Tulane Flowerree Symposia on Cognition, Psychology Press, 1989, pp. 107–133.

[115] J. G. TRAFTON, N. L. CASSIMATIS, M. D. BUGAJSKA, D. P. BROCK, F. E. MINTZ, AND A. C. SCHULTZ, *Enabling effective human-robot interaction using perspective-taking in robots*, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 35 (2005), pp. 460–470.

[116] J. G. TRAFTON, L. M. HIATT, A. M. HARRISON, F. P. TAMBORELLO, S. S. KHEMLANI, AND A. C. SCHULTZ, *Act-r/e: An embodied cognitive architecture for human-robot interaction*, Journal of Human-Robot Interaction, 2 (2013), pp. 30–55.

[117] M. A. UMILTA, E. KOHLER, V. GALLESE, L. FOGASSI, L. FADIGA, C. KEYSERS, AND G. RIZZOLATTI, *I know what you are doing: A neurophysiological study*, Neuron, 31 (2001), pp. 155–165.

[118] D. Vanderelst and A. Winfield, *An architecture for ethical robots inspired by the simulation theory of cognition*, Cognitive Systems Research, 48 (2018), pp. 56–66.

[119] A. Vinciarelli, A. Esposito, E. André, F. Bonin, M. Chetouani, J. F. Cohn, M. Cristani, F. Fuhrmann, E. Gilmartin, Z. Hammal, D. Heylen, R. Kaiser, M. Koutsombogera, A. Potamianos, S. Renals, G. Riccardi, and A. A. Salah, *Open challenges in modelling, analysis and synthesis of human behaviour in human–human and human–machine interactions*, Cognitive Computation, 7 (2015), pp. 397–413.

[120] A. Wagner and E. Briscoe, *Psychological modelling of humans by assistive robots*, (2016), pp. 273–296.

[121] J. Wainer, D. J. Feil-seifer, D. A. Shell, and M. J. Mataric, *The role of physical embodiment in human-robot interaction*, in ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication, 2006, pp. 117–122.

[122] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric, *Embodiment and human-robot interaction: A task-based perspective*, in RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication, 2007, pp. 872–877.

[123] Y. Wang and G. Ruhe, *The cognitive process of decision making*, (2007).

[124] Y.-H. Wang, G.-Y. Liu, G. Huang, and Y. Wang, *Variable admittance force feedback device and its human-robot interaction stability*, Robotics and Computer-Integrated Manufacturing, 82 (2023), p. 102537.

[125] Z. Wang, *Modelling decision-making in a joint action for picking an object*, Master's thesis, University of Bristol and University of West of England, 2020.

[126] C. Wilimzig, S. Schneider, and G. Schöner, *The time course of saccadic decision making: Dynamic field theory*, Neural Networks, 19 (2006), pp. 1059–1074.

[127] A. F. Winfield, *Robots with internal models: a route to self-aware and hence safer robots*, (2014).

[128] D. M. Wolpert, K. Doya, and M. Kawato, *A unifying computational framework for motor control and social interaction*, Philosophical Transactions of the Royal Society of London B: Biological Sciences, 358 (2003), pp. 593–602.

[129] J. D. WONG, T. CLUFF, AND A. D. KUO, *The energetic basis for smooth human arm movements*, Elife, 10 (2021), p. e68013.

[130] R. E. WRAY AND R. M. JONES, *An introduction to soar as an agent architecture*, Cognition and multi-agent interaction: From cognitive modeling to social simulation, (2005), pp. 53–78.

[131] ——, *Considering soar as an agent architecture*, Cognition and multi-agent interaction: From cognitive modeling to social simulation, 33 (2006), pp. 53–78.

[132] A. YAMAZAKI, K. YAMAZAKI, Y. KUNO, M. BURDELSKI, M. KAWASHIMA, AND H. KUZUOKA, *Precision timing in human-robot interaction: coordination of head movement and utterance*, in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2008, pp. 131–140.

[133] J. YEN, J. YIN, T. R. IOERGER, M. S. MILLER, D. XU, AND R. A. VOLZ, *Cast: Collaborative agents for simulating teamwork*, in International Joint Conference on Artificial Intelligence, vol. 17, LAWRENCE ERLBAUM ASSOCIATES LTD, 2001, pp. 1135–1144.

[134] X. YU, S. ZHANG, Y. LIU, B. LI, Y. MA, AND G. MIN, *Co-carrying an object by robot in cooperation with humans using visual and force sensing*, Philosophical Transactions of the Royal Society A, 379 (2021), p. 20200373.

[135] Z. ZHANG, N. RUSSWINKEL, AND S. PREZENSKI, *Modeling individual strategies in dynamic decision-making with act-r: A task toward decision-making assistance in hci*, Procedia computer science, 145 (2018), pp. 668–674.

[136] C. ZHU, Q. CHENG, AND W. SHENG, *Human intention recognition in smart assisted living systems using a hierarchical hidden markov model*, in 2008 IEEE International Conference on Automation Science and Engineering, 2008, pp. 253–258.