Towards Improving Human-Robot Interaction For Social Robots

Saad Khan
University of Central Florida

Towards Improving Human-Robot Interaction For Social Robots

by

SAAD AHMAD KHAN
B.S. Electrical Engineering, University of Engineering & Technology, Lahore 2007
M.S. Computer Engineering, University of Central Florida, 2013

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Electrical and Computer Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2015

Major Professor: Ladislau Bölöni

ABSTRACT

Autonomous robots interacting with humans in a social setting must consider the social-cultural environment when pursuing their objectives. Thus the social robot must perceive and understand the social cultural environment in order to be able to explain and predict the actions of its human interaction partners. This dissertation contributes to the emerging field of human-robot interaction for social robots in the following ways:

1. We used the social calculus technique based on culture sanctioned social metrics (CSSMs) to quantify, analyze and predict the behavior of the robot, human soldiers and the public perception in the Market Patrol peacekeeping scenario. 2. We validated the results of the Market Patrol scenario by comparing the predicted values with the judgment of a large group of human observers cognizant of the modeled culture. 3. We modeled the movement of a socially aware mobile robot in a dense crowds, using the concept of a *micro-conflict* to represent the challenge of giving or not giving way to pedestrians. 4. We developed an approach for the robot behavior in micro-conflicts based on the psychological observation that human opponents will use a consistent strategy. For this, the mobile robot classifies the opponent strategy reflected by the personality and social status of the person and chooses an appropriate counter-strategy that takes into account the urgency of the robots' mission. 5. We developed an alternative approach for the resolution of micro-conflicts based on the

imitation of the behavior of the human agent. This approach aims to make the behavior of an autonomous robot closely resemble that of a remotely operated one.

بِسْمِ اللهِ الرَّحْمَنِ الرَّحِيمِ

In the name of Allah, the Beneficent, the Merciful.

This work is dedicated to the most wonderful parents in the world, Zubair Ahmad Khan and Farakhanda Akhtar whose advice and support helped the most throughout the entire doctorate program. You have helped me through my most stressful times and all the hard work has paid off. I can't begin to thank you for all you have done over the years for me. Your sacrifices, support and encouragement made it possible.

ACKNOWLEDGMENTS

I am most grateful to the members of my committee, Dr. Steve Fiore, Dr. Gita Sukthankar, Dr. Aman Behal, and Dr. Ivan Garibay for their time, encouragement, and expertise throughout this work.

Special thanks to the chairman of my committee, Dr. Ladislau Bölöni, for his exquisite attention to detail and for his demand for excellence. He has been an amazing mentor and I'm extremely grateful for his ongoing feedback, support and consistent push for critical thinking that challenged me throughout this work.

There are people in everyones lives who make success both possible and rewarding. A special thanks and feeling of gratitude towards my loving siblings, Fahad Ahmad Khan and Sana Ahmad who have never left my side and are very special.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

Social robots must obey the social and cultural norms of the environment in which they operate. Besides obeying these rules, the robots should also be aware of their mission goals. Except in the specific case when the only goal of the robot is to behave in a socially acceptable way, there can be a conflict of interest between the objective of the robot to accomplish its mission and the constraints of the social and cultural norms. Achieving appropriate behavior in a social-cultural context is one of the most elusive goals of agent research. There are, however, many practical applications where social behavior is necessary. Agents acting in virtual environments, such as games or training must show a believable social behavior. This can often be achieved with careful scripting. However, when agents control autonomous robots that interact with humans in social settings, the requirements are harder and the interactions more open ended. The agent must have a model to evaluate the impact of specific actions on the participants in the social interaction. There are actions which are physically possible, but socially unacceptable in a given culture. We will use the term *social calculus* for this evaluation process.

The fields of sociology and psychology have a rich literature of describing human behavior in specific cultural contexts. Social calculus, however, requires explicit formulas or algorithms that take as input the observable facts of a situation and specific actions, and provide an output in the form of quantitative metrics. The models developed in humanities are rarely expressed in such quantitative form. In recent years, there is an ongoing effort to *operationalize* models from sociology and psychology [1, 2, 3]. Alternatively, we can design new models of reasoning in a social-cultural context, which are informed by the sociological models, but designed from ground up to provide an implementable algorithmic framework.

In Chapter 3.1, we describe the framework of culture-sanctioned social metrics (CSSM) [4, 5, 6]. We assume that the human behavior proceeds through a series of actions a_i. Actions impact the state of the actor, the target of the action, their peers as well as the perception of the general public. In this model, the state of the agent, relevant to its actions in the social-cultural context is described by a collection of metrics. The metrics can be divided into tangibles (such as wealth and time) and the socially constructed CSSMs (such as dignity and politeness). CSSMs are not necessarily independent, but they are not arbitrarily convertible to each other. For modeling CSSMs in a real-world scenario, we consider a cross-cultural social interaction scenario from a type reportedly encountered peacekeeping missions. The scenario describes a series of social interactions between a soldier and a local merchant at a military checkpoint located at the entrance of a busy marketplace. The scenario illustrates the complex balance between mission objectives, cultural sensitivity and trust-building actions.

In order apply the model to a given scenario involving one or more cultures, we need to (a) choose a set of CSSMs appropriate to the culture and (b) acquire the action-impact functions for all the feasible action combinations.

Part (a) of the task is clearly a task for a social anthropologist. CSSMs are strongly tied to human cultures: they cannot be inferred from first principles. The translation of the name of a CSSM into a foreign language or its use in a different cultural context might not transfer the evaluation algorithms and rules of conduct. For instance, the term “dignity” has different evaluation methods and rules of conduct in the African-American culture compared other English-speaking cultures. The sociological concept of “face” has three different words in Chinese: *mian*, *lian* and *yan* [7]. The relatively well established terms of “loosing face” and “saving face” are Chinese lexical borrowings, which entered English in the late 19th century. In other languages, such as Hungarian, these concepts can be explained only through circumlocutions. In Chapter 4, we validate the selection of the CSSMs for the Middle East marketplace scernario.

For part (b), the task of designing the AIFs, the situation is different. As AIFs are multiparameter mathematical functions, we cannot directly ask them from human informants. Knowledge engineering these functions for every possible action is a difficult challenge, because the design space is very large. In [6] we have modeled the Spanish Steps flower selling scam using the CSSM mechanism. The scenario has only two participants - yet there are 20 different actions, 14 different CSSMs (if we consider self, peer and public perceptions separately). This is already a significant knowledge engineering task. As we are moving to more open-ended scenarios, with a larger number of participants, the number of AIFs and their respective complexity increases at least quadratically. Finding efficient methods to acquire the AIFs is thus a critical step in making the CSSM approach applicable to medium size real world interaction scenarios.

In Chapter 4, we describe a method to acquire AIFs from a survey of human respondents for specific spot values of the actions. Using these inputs, we evolve the AIF functions using a genetic programming mechanism. The objective is to find functions that match the input, provide realistic results when becoming part of the agents and can be expressed in mathematically simple forms. We also hope that the mathematical form of the evolved functions will have explanatory power about the human social-cultural behavior in the given context.

In Chapter 5, we focus on human social behavior for the movement in crowds. For our environment, we consider a busy marketplace where physical obstacles are combined with crowds of people. The individual members of the crowd behave in a purposeful way: move from one shop to another, stop at various landmarks or head towards the exit along a pre-planned but not rigidly fixed trajectory. We will say that the individuals have a *mission* with a specific value and urgency. The movement of people in such environments is governed by social norms: they are not supposed to violate each other's personal space, block each other's intended direction of movement or physically bump or push each other. The social norms for physical movement depend on the culture and social setting. Different cultures define the personal space of an individual differently, and put different penalty on physical contact. Whether movement in a certain environment can be performed without violating any social norm depends on the density of the crowd: beyond a certain density, an individual who tries to avoid any violation of personal space will not make any progress at all. Groups of individuals moving in dense crowds will enter into *micro-conflicts* if following their planned trajectory would create an unplanned, large social cost through physical collision or severe violation of personal space. The attribute *micro* illustrates the fact that these conflicts are normally resolved in several seconds: one or more participants will alter their speed and/or path, reducing the social cost to an acceptable level.

To successfully participate in micro-conflicts, the robot needs to have an operational model of the social costs and mission costs as perceived by the local society and the individual participants. Furthermore, it needs to have a model of the strategies deployed by the human participants when participating in the micro-conflicts, such that it can develop effective counter-strategies that either mimic or extend those used by humans. In Chapter 6, we study the learning of a consistent micro-conflict strategy against different types of strategies which humans might deploy in such scenarios. We are particularly interested in how a *consistent* strategy would look like in this setting, and how a robot might emulate this. One of the important insights is that as important as it might seem to obey all the social rules, in a sufficiently dense crowd it is impossible to completely avoid incurring any social cost.

Micro-conflicts can be modeled as a two-player two-move game, however, the strategy played by the participants are not necessarily optimal in a game theoretic sense. The human players are influenced by external factors, personal history, psychology and so on. It is possible that a human player will concede the right of way to four passersby but, annoyed by the wasted time, he will aggressively cut off in front of the fifth. The objective of studying the features of human behavior while navigating in a crowd, and to develop the techniques through which human strategies can be learned by imitation has been explored in Chapter 7. We are especially interested in finding out whether human behavior is consistent in the game theoretic sense, and if not, what other factors besides the game payoffs might affect the behavior.

CHAPTER 2
RELATED RESEARCH WORK

2.1 Models of Social and Cultural Behavior

We model the calculus of human social-cultural behavior with the objective to provide explanatory and predictive power. We will discuss related work in two different fields:

- Models of social and cultural behavior in the social sciences - such as psychology, sociology and anthropology. Although these fields favor the form of a narrative rather than formal description, many researchers have expressed their insights in a numerical form, which can be relatively easily translated into computational models.

- Models of human social behavior in engineering, built with the goal of a specific application. Engineering solutions are often based on formal models or inspired by theories developed in social sciences. Nevertheless, the practical requirements of an engineering problem, such as the scarcity of available data and performance considerations sometimes led researchers to start from a blank state, and build problem-specific models based on purely engineering considerations.

2.1.1 Social and Cultural Models in the Social Sciences

One of the most influential models from our perspective is politeness theory, initiated by Brown and Levinson [8], and extended by many other researchers. The overall assumption is that politeness centers around the maintenance of "face" defined as the public self image of the adult human. More specifically, they define the "positive face" which refers to one's self esteem and the "negative face" which refers to one's freedom to act.

The Brown and Levinson model is often interpreted in terms of the work of Paul Grice [9] who formulated the *cooperative principle* in conversations. According to the four maxims formulated by Grice speakers in a collaborative conversation should be truthful, provide an appropriate amount of information (not too much, not too little), be relevant and avoid obscurity of expression.

Almost always, the desire to be polite (in the Brown and Levinson definition) and the desire to be cooperative (in the sense of Grice's maxims) are countervailing forces. For instance, the indirect strategy is highly polite, but leads to inefficient communication.

The Brown-Levinson model, by positing two metrics which humans want to maximize, was one of the direct influences for our approach of defining CSSMs. The most significant difference is that CSSMs are easy to collect: the intra-cultural uniformity conjecture implies that we can ask any member of the culture to evaluate them. In contrast, the terms positive and negative face do not mean anything to an untrained participant; their values must be evaluated by people with significant training. Furthermore, both the Brown-Levinson and

Grice models attempt to discover the culture-independent universals in human communication. The Brown-Levinson definition of politeness does not necessarily match the definition of politeness and indeed the desirable behavior in specific cultures. There are cultures, for instance, where direct speech is considered polite and desirable. The interpretation of the Brown-Levinson model in the context of specific cultures is a significant ongoing research topic [10, 11].

Another influential model, which specifically attempts to account for and quantitatively measure cultural differences, is the cultural dimensions theory of Geert Hofstede [12]. In the most recent publications, six dimensions are considered: (1) power distance, the acceptance of unequal distribution of power, (2) individualism versus collectivism, (3) uncertainty avoidance (4) masculinity versus femininity, a metric measuring the balance between assertiveness and competitiveness versus a focus on cooperation, human relations and quality of life, (5) long term versus short term orientation and (6) indulgence versus self-restraint. From the point of view of our model, CSSMs can be associated with one or more of these dimensions - for instance dignity has relevance to (1) and (4), while wealth to (5) and (6). Furthermore, Hofstede's analysis shows us that even if two cultures define the same set of CSSMs, they might weight these CSSMs differently in practical behavior.

2.1.2 Social and Cultural models in Engineering

Social models developed in engineering aim to develop artifacts such as software agents, avatars, websites or robots which take into account the social and cultural environment in which they are used. Requirements of practical applicability dominate in these fields. Engineering artifacts face additional challenges in their deployment, for instance the problem of sensing the social signals made by humans (see Vinciarelli et al. [13]).

Some of the engineering research is directed explicitly towards the practical deployment of models proposed in the social sciences.

Miller et al. [1] describes a software product called the Etiquette Engine which uses the Brown-Levinson politeness model [8] to assess the politeness in interactions involving military personnel of common culture but different rank (such as the interaction between a corporal and a major). In a follow-up work [2] the authors create a more complex model which investigates how culture (as examplified by Hofstede's cultural factors [12]) as well as politeness levels affect the way in which people react to instructions, commands or requests ("directive compliance").

Bosse et al. [3] formalizes Damasio's theory of consciousness [14], where consciousness is built up from the distinct elements of emotion, feeling and core consciousness, the latter being defined as the "feeling of a feeling". The authors use the model of state properties described as large-multi dimensional vectors. The dynamics of these models is described using a logic-based temporal trace language (TTL). The authors verify through simulation that the model indeed exhibits the properties posited by Damasio's model.

The ubiquity of user interfaces featuring synthetic characters naturally led to the requirement that they exhibit appropriate social and cultural behaviors, such as empathy. Paiva et al. [15, 16] describe the functionality of a virtual environment called "Fear not!" which allows 8-12 year old children to witness bullying situations from a third person perspective. To model the emotions of the characters in the simulation, the system uses the cognitive theory of emotions of Ortony et al. [17]. An extension of this work is described by Rodrigues et al. [18] where the empathy model relies on the neuropsychological theories of Perceptual Action Model (PAM) [19] and the work of Vignemont and Singer on the emphatic brain [20].

Another relevant point of view is that of *social intelligence*, defined as the ability to act for social benefits. For instance, Hogg and Jennings [21] describes a model for socially intelligent reasoning for autonomous agents. The authors rely on Harsanyi's social welfare function [22] to balance the benefits to others in the course of taking an action and weight it against its own benefits.

As mobile robots are increasingly deployed in situations with human interaction partners and bystanders [23], the field of human-robot interaction [24] must increasingly consider issues of social intelligence [25].

The Kismet robot (Breazeal [26]) was developed in the context of the Sociable Machines Project at MIT. The principles behind the robot integrated theories of infant social development, psychology, etology and evolution. The robot was able to infer emotions in the human users, and to emulate and display emotional states such as anger, fear, disgust or sorrow.

While we can learn many practical lessons from sociable robot projects, there are also several important differences. Inevitably, these projects put the robot front and center, and the social interactions are always modeled between the robot itself and the interaction partner. In contrast, our work models general purpose social interactions, with or without the active participation of robots. Another major difference is that projects such as Kismet consider theories of emotions which are actually felt by the human user (but are only emulated by the robot). In contrast, CSSMs are imposed from outside, by the culture. While Kismet models the ways humans are guided by their *emotions*, we are modeling behaviors guided by social and cultural *conventions*.

2.2 Mobile Robot Movement in Dense Crowds

Part of the contributions of this dissertation considers the movement of a social robot in a dense crowd. There is a diverse research literature dealing with the movements of robots in crowds. Naturally, for any such work the first step is to develop an understanding of crowd behavior, which is dependent of situations, environment and of course social and cultural factors.

The minimal goal for a mobile robot acting in a crowd is to avoid robot-human collisions, for instance using obstacle avoidance technologies based on either velocity vectors or potential fields.

In contrast to purely physics based approaches, our approach to this problem is based on a game-theoretic model of human-robot interaction in the micro-conflicts, a work that must be seen in the greater conflict of game-theoretic modeling of human adversaries. Finally, one of our contributions is an approach that sees the objective of the robots is to behave as if it is under the control of a human operator - a work that must be positioned into the imitation learning literature.

2.2.1 Modeling Crowd Behavior

Social behaviors are learnt from social interactions and evolve over the course of multiple social interactions [5, 4]. For collision avoidance in crowds, pedestrians would follow those movement conventions that would easily allow them to avoid collisions. For instance, a person from Europe would prefer to step aside on the right side for avoiding collision. In South Asia, people move towards their left to avoid collisions [27]. Nevertheless, a pedestrian always has the free will to select any movement in crowd (exceptions exist - a campaign by the government in South Korea urges people to walk on the right-hand side [28, 27]). But there are certain actions that are not socially acceptable, e.g., deliberate transgression of pedestrian's personal space.

Human crowds had been modeled through a wide range of techniques. Social force based models [29, 30] had been found especially useful in modeling emergency situations where dense crowds are acting in an unplanned manner (the “fire in the theater” scenario). We can see social force models as modeling the instinctive aspects of human behavior. In contrast, when the crowd members have more time to make conscious decisions about their actions, the social conditioning aspects of crowd behavior become more important [31]. More recent approaches use a combination of models including psychological and geometrical rules as well as social and physical forces [32].

One of the important questions with regards to crowd dynamics is the presence of a mobile robot in the crowd. There are two new aspects to consider. One of them is how the presence of the crowd influences the movement of the robot – for instance, there is a possibility that a cautious robot will freeze up in a dense crowd [33]. Another aspect, possibly highly relevant in the future is how the presence of the robot modifies the behavior of the crowd, and whether this impact can be exploited for crowd control [34].

2.2.2 Velocity Obstacle Avoidance Methodologies

Whenever we are considering the control of multiple mobile agents (robots, vehicles or airplanes), the avoidance of collisions is one of the most significant challenges. The collision avoidance problem can posed as geometric optimization problem which takes into account the number of static and dynamic obstacles. There are several approaches through which such optimization problems can be solved, for instance through linear integer programming, or geometric approaches such velocity obstacle (VO) avoidance techniques for local collision avoidance. [35] uses the concept of a collision cone as the basic geometric shape for collision avoidance. [36] selects its optimal velocity from the set of permitted velocities using linear integer programming. The set of permitted velocities is selected with respect to the geometric space of velocity obstacles induced by other moving agents. The inherent problem faced by using such techniques is the oscillating velocities while avoiding other agents. This means that if two agents are using a similar collision avoidance technique, i.e., if they are selecting new velocities outside the pool of velocity obstacle induced by the other agent then their old velocities will become a part of velocity obstacle for new selected velocities. Hence, agents will move back to their previous velocity and the agents will oscillate back and forth in a region of permitted velocities. To counter this problem, [37] introduced the technique of RVO (reciprocal velocity obstacle avoidance) which shares the responsibility of avoiding collision with the opponent agent. OCRA [38] further extends the concept of RVO considering n-agents path planning using the geometric optimization technique.

2.2.3 Potential Field Methods for Collision Avoidance

Another well known approach for collision avoidance in the multiagent systems is the potential field method [39, 40]. The basic concept is the use of artificial potential fields inside the workspace of a robot that is attracted towards its goal and repelled by the obstacles.

The workspace is discretized into a regular grid and each cell corresponds to the sum of the repulsive potential generated by obstacles and attractive potential generated by the goal position. Therefore, gradient methodologies are applied to maneuver with collision avoidance towards goal. Gradient techniques are prone to the problem of local minima[41]. One of the solutions to this problem is to utilize potential fields that are solutions to the Laplace equation (harmonic functions) [42].

2.2.4 Game Theoretic Model for Human Adversaries

An interesting class of game theoretic approaches governing encounters between mobile agents are based on modeling the human adversaries using Stackelberg games. Most of these approaches consider a patrolling strategy, where the goal is not the avoidance of a collision, rather the facilitation of patrolling, where opponent agents actively try to avoid the patrol [43, 44, 45]. This hide and seek game can be modeled as the zero-sum strategic game where the hider selects the cell from the grid, and the seeker seeks (selects) the cell chosen by the hider. Modeling in terms of a Stackelberg game with repeated interactions, the strategy selection by follower (hider) is assumed to be optimal based on the leader's (seeker) strategy. The possibility for the hider to observe the seeker's strategy before committing to its own strategy radically influences the outcome of the game. But as humans deviate from optimal selection due to irrational behavior, it is necessary for the leader to incorporate such irrational behavior in its strategic model. In [46] three such algorithms are introduced, based on mixed integer linear programming which effectively handles the uncertainties due to bounded rationality and limited observations of adversary. Some of these algorithms are currently being actively deployed (GUARDS[47], PROTECT[48]).

2.2.5 Imitation Learning in Crowds

The dissertation work fits in a larger trend in robotic systems

Learning is an important technique in designing robot behaviors as the explicit programming of robots takes requires significant expertise and effort. Reinforcement learning allows a robot to teach itself by trial and error. From the robot's perspective, reinforcement learning is a way to find a policy which optimizes a function - which must be clearly specified and provided to the robot. The reinforcement learning robot also needs a way to evaluate the state of the system for rewards and penalties. We can contrast with this a learning technique often denoted with terms such as learning from demonstration, imitation learning or learning by showing [49, 50]. From the perspective of such an optimization, learning by demonstration can be seen as a shortcut through the experimentation required by reinforcement learning by having a teacher provide a demonstration of the desired actions in selected scenarios - which the robot can use to learn and refine its own behavior. It had been shown that learning by demonstration can create policies with higher performance than that of the teacher. In addition, the learning might proceed through social interaction and partnership, with no clearly defined roles for the teacher and student [51].

Note, however, that learning by demonstration is also possible when the optimization goal is not explicitly specified. In this case, the robot will imitate salient features of the behavior of the teacher, without having a good model to understand why the teacher does what it does. Albeit this might seem strange, human and animal imitation often proceeds like this [52, 53], especially with regards to social behavior. We do not teach children a theory of social behavior, but illustrate what actions must be taken in specific circumstances. Some researchers had proposed restricting the term of imitation to this kind of scenarios [54].

One possible classification of the learning by demonstration robotics systems is the mechanism of actual learning. Such a system can learn (a) directly a policy which maps from states to actions, (b) a dynamical model of the system from which a policy can be later extracted or (c) a causal model of the system which associates pre- and post-conditions to actions, from which a planner can create a policy. Our approach, fits in the first category.

Another way to classify such systems is the choice of demonstrator body: in contrast to systems where the demonstrator is using its own body to demonstrate the desired behavior, in our case, the demonstrator body is actually that of the robot, which is remotely controlled by the teacher. This fits well with the projected use in a mixed autonomy system, and simplifies the *embodiment mapping* component of the system (which is an identity mapping). We need to mention that it is possible to learn from demonstrators with a different body architecture (for instance, a bipedal robot learning from a human [55]), even from a heterogenous mix of demonstrators [56, 57].

The system still needs to deal with the challenge of *record mapping* which maps the environment experienced by the teacher to that of the learner. In our case this mapping happens through the choice of features, most of them related to the micro-conflict games.

CHAPTER 3
SOCIAL CALCULUS - THE CHECKPOINTSCENARIO

CSSMs are consistent in a given culture, but they vary between cultures. A given culture assigns a name, a calculation method and a series of behavior rules to these metrics. Agents not immersed in a particular culture would not know about, or would not know how to calculate these values. Even an agent which is immersed in the culture might choose to ignore the rules associated with these values (but it would be aware of the transgression). Finally, an agent might not be able to accurately observe or compute the values (which frequently require a significant cognitive load and accurate observation of the environment). Agents might also make mistakes when planning their actions - especially in cases when they interact with agents which use a different set of values. The latter cases constitute cases of bounded rationality.

3.1 The Marketplace Checkpoint Model

For modeling the real world scenario, we used the running example which is a situation frequently encountered in peacekeeping missions. The scenario is the series of social interactions outside a military checkpoint that is located at the entrance of a busy market. We assume the location to be a Middle Eastern country (although the scenarios would unfold roughly similarly in other parts of the world - with the necessary adaptations for the cultural specifics). The checkpoint is manned by a sergeant (S), a private (P) and a robot (R). A street vendor (V) takes advantage of the traffic slowdown by positioning its cart near the checkpoint at one of the four feasible locations L1-L4 (see Figure 3.1, at increasing distance from the checkpoint, our modeling will be concerned with the interactions between these actors over the course of several weeks. Let us now informally describe the various values, considerations and possible actions which are at stake at this scenario.

Figure 3.1: The private P is interacting with vendor V, with the sergeant S and robot R in the background.

Soldiers on peacekeeping missions need to balance their own security and military objectives with the need to maintain a friendly relationship with the local population. Our work is an attempt a quantitative, operational model of the ways in which various actions taken by the soldiers (and in the near future, robots) as well as the members of the local population impact their respective cultural values and perceptios of each other. Some of the obvious challenges in this work include:

- The difficulty to assign numerical metrics and calculations to values dependent on social, cultural and personal perception.
- The need to consider the interaction between multiple players, some of them individual (soldiers, members of the local population, the robot) but some of them groups of people (e.g. the participants in a crowd).
- The need to consider the evolution of values over a longer amount of time. The evolution of certain values, such as gaining of trust can not take place over a single interaction. On the other hand, single interactions must be considered, as certain gestures might have a long lasting impact.

Although the literature on cultural interactions is vast, most of the research done in the humanities do not generate an *operational* model. Even when explicit numerical values are given (such as in Hofstede's models [12]) the values are averaged over the general population, and they can not be used to characterize individual behavior.

In contrast our objective is to develop a system that allows automated analysis of a specific scenario, with actors who are members of their respective cultures, but are also identifiable individuals with a high degree of freedom in their choice of actions.

Such an system can be used immediately as a training or assessment tool. It can also serve as a modeling tool to aid policy making, and, in the future, a component of the robot behavior agent.

The POV of the checkpoint team: the efficiency of the checkpoint and their personal security require maintaining a free and uncluttered area around the checkpoint. On days with a high alert level the perceived security is lower, and due to the more thorough inspections the traffic through the checkpoint slows down. The presence and location of the food vendor affects the security risks. Security threats can come from the street vendor itself, from creating additional crowding near the checkpoint, and from blocking lines of sight (either directly, or through the crowding).

The checkpoint team considers desirable to maintain good relations with the local population (in general), and the food vendor (in particular). Friendly interaction (informal conversations, exchange of gifts) increase friendship and trust. Unfriendly actions (such as ordering around or threatening) negatively impact the relations.

The POV of the street vendor: it is in the financial interest of the vendor to position its cart closely to the checkpoint. He will try to maintain friendly relations with the members of the checkpoint team, and will remember past interactions with the individual soldiers, appropriately reciprocating friendly or unfriendly behavior. He is aware of factors such as high alarm level (which can mitigate a specific intransigence from the checkpoint team). On the other hand, impolite behavior from a soldier which is considered a friend is perceived more negatively than, for instance, impolite behavior from the robot. The vendor will follow his cultural norms in his behavior - for instance, it is not acceptable to refuse a polite request from a friend.

3.2 CSSM for the Checkpoint Scenario

3.2.1 The Choice of CSSMs

Let us now analyse and model out scenario using the CSSM model. We shall use the following collection of metrics:

- **Financial worth (V):** the income of the seller. It is dependent on the location, scaled by the traffic of the given day, and limited by the maximum amount of clients the seller can handle. It is measured in the local currency. It is only relevant to the client.
- **Perceived security level (S, P, R):** is a metric of the level of threat as perceived by the soldiers. It depends on the alarm level, on the level of traffic, and the crowd created by the vendor.
- **Dignity (S, P, V)**. The perception of the personal dignity by the soldiers and the vendor, for the sake of simplicity we shall call both of them dignity, but the two parties apply different evaluation algorithms. The soldiers use a generic Western cultural model adapted to their status as soldiers (being defied on an open order decreases dignity). The seller uses its own cultural model - for the actions of this scenario, for

instance involves that being ordered around decreases dignity. Similarly the refusal of an offered gift is an offense to the vendor.

- **Politeness (S,P, V)** The perceived politeness metric is evaluated according to culture specific algorithms by the vendor and the soldiers.

3.2.2 Action Repertoire

We model the possible scenarios using a series of possible actions. An action is performed either by a single actor (e.g. the vendor V moving from L1 to L3) or is the interaction between an actor and a recipient (the vendor V giving a gift to sergeant S). From the point of view of our model, the actions are fully described by their impact on the values of the actor and (if applicable) the recipient. Our modeling approach here is to define a relatively small number of actions, but to characterize them with detail variables which describe, for instance, the destination of a movement or the verbal style in which a request or command is delivered. These actions are listed in Table 3.1.

3.2.3 Case Study of an AIF

One of the most critical and interesting actions is A6, where the the representative of the soldiers (S, P or R) requests the vendor V to move the cart to a farther location. This

Table 3.1: Possible actions for the participants in the Market Patrol scenario (with specific possibilities for actor and target)

Action		Actors	Targets	Parameters
A1	moves	V		Location
A2	declines-to-search	V		Offensiveness
A3	offers-gift	V	S, P	
A4	initiates-conversation	V, S, P	V, S, P	
A5	accepts-conversation	V, S, P		
A6	orders-to-search	S, P, R	V	Offensiveness
A7	passes-order	S, P	P, R	
A8	accepts-gift	S, P	V	
A9	declines-gift	S, P	V	Offensiveness
A10	order-to-move	S, P, R	V	Loudness
A11	overnight	S, P, R, V		

requests goes against the financial interests of the vendor. What we need to investigate is how this request (and the response to it) affect the values of the participants.

First of all, we need to discuss the detail variables of the action A6. This request can be made at various levels of politeness. To find a numerical metric of the politeness level of a request, we will use the *mitigation level* of the order - according to the classification recently popularized by Malcolm Gladwell [58][1]. To the six mitigation levels discussed by Gladwell, which culminate in command, we add three more levels which model the threat of and actual physical actions, respectively.

Note that the robot is not expected to know the subtleties of polite conversation, thus its use of direct command mode carries less offense - and its own politeness is irrelevant and not measured.

This fact opens interesting possibilities for action strategies from the point of view of the team.

Note that the values in the table are calculated from a Middle Eastern perspective. Certain cultures such as Korean or Japanese, would put a significantly higher penalty on unmitigated speech. On the other hand, Northern European cultures would not put virtually any penalty on direct speech (and high level of mitigation would probably be incomprehensible).

Similar considerations apply for the action of the refusal by the vendor to move to the suggested location (which can be also be done with different levels of mitigation).

The values in the table can also be modeled in an equation form using a combination of signum, heaviside, exponential and other simple mathematical functions:

[1]Note however, that similar ideas are present in the literature for a long time - e.g. in Brown and Levinson's politeness model[8]

$$\begin{aligned} F(s5,a6)_{s,p} &= sgn(5-x)\left[|5-x| + \frac{sgn(5-x)}{y+z}\right] \\ F(s3,a6)_v &= -H(x-4)\cdot e^{x/3} \end{aligned} \tag{3.1}$$

where, x is the level of mitigated speech, y and z are the loudness and offensiveness respectively. In Equation 3.1, the function *sgn* is the signum function, whereas *H(x)* is the Heaviside's function. In Chapter 4.1, we provide the genetic learning procedure through which one is able to formulate CSSM AIF's.

Table 3.2: The impact of action A6 on the politeness of soldiers S or P and the dignity of the vendor using various levels of mitigated speech

Name	Example	$P^{S/P}$	D^V
L1: Hint	Seems like you have got new stuff in your bag to sell in market today.	1.0	1.0
L2: Preference	I like the stuff you sell, and would love to share my opinion about your new items (in the bag)	0.81	1.0
L3: Query	Won't you show me the new stuff that you're going to sell today?	0.68	1.0
L4: Suggestion	I would suggest that you let me search the bag, as the security alert is high today	0.56	0.91
L5: Obligation statement	I'm sorry i need to do this, but my boss insists that you show me your bag	0.44	0.73
L6: Command	Show me your bag!	0.36	0.63
L7: Threat of physical action	Show me your bag or i'll have to arrest you!	0.22	0.49
L8: Minor physical action	Pushing and snatching the bag, afterwards going through bag without consent of vendor	0.11	0.28
L9: Major physical action	Taking the vendor in custody	0	0

Table 3.3: The impact of actions on the values of the vendor and the soldiers

The actions of the solider

The social values of the vendor

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
S1	F(s1, a1)$_v$	0	0	0	0	0	0	0	0	0
S2	0	0	0	0	0	0	0	0	0	0
S3	0	0	0	0	0	F(s3, a6)$_v$	0	10	F(s3, a9)$_v$	0
S4	10	F(s4, a2)$_v$	15	5	5	0	0	0	0	0
S5	F(s5, a1)$_v$	F(s5, a2)$_v$	10	F(s5, a4)$_v$	5	F(s5, a6)$_v$	0	0	0	0
S6	0	-10	0	0	F(s6, a5)$_v$	0	0	0	0	-15
S7	5	-20	0	0	0	0	0	0	0	0
S8	0	-10	0	0	F(s8, a5)$_v$	F(s8, a6)$_v$	0	0	-5	0

Table 3.4: The impact of actions on the values of the sergant and private

The actions of the solider

The social values of the sergant and private

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
S1	0	0	0	0	0	0	0	0	0	0
S2	F(s2, a1)$_{s,p}$	F(s2, a2)$_{s,p}$	0	0	0	0	0	0	0	0
S3	0	F(s3, a2)$_{s,p}$	0	0	0	0	0	0	0	0
S4	0	0	0	5	5	F(s4, a6)$_{s,p}$	0	10	F(s4, a9)$_{s,p}$	-20
S5	0	-10	0	F(s5, a4)$_{s,p}$	5	F(s5, a6)$_{s,p}$	0	5	0	0
S6	F(s6, a1)$_{s,p}$	-10	5	0	F(s6, a5)$_{s,p}$	0	0	10	0	0
S7	10	-10	0	0	0	0	0	0	0	0
S8	0	-10	0	0	F(s8, a5)$_{s,p}$	0	-10	10	F(s8, a9)$_{s,p}$	-20

3.2.4 Beliefs and Public Perception

The impact of an action on a culture sanctioned value is modulated by the beliefs of the agent about specific aspects of the current context. A culture requires its members to maintain these beliefs as accurate as possible - the correctness of beliefs are necessary for the culture to operate as expected. Nevertheless, it is quite possible for an agent to have incorrect beliefs, especially in inter-cultural exchanges, when the agent might mis-interpret the social signals (computers are especially bad at this, see Vinciarelli et al. [13]). Even when incorrect, beliefs are important, because the agents will act and calculate CSSMs according to the beliefs, whether they are correct or not. If an agent considers another one a friend, it will act accordingly and judge the actions of the other agent in this context, irregardless if the friendship is mutual or not.

In the agent literature, the beliefs of the agent are frequently considered to be a "model of the world". Creating such a model, for human players, is clearly impossible. We argue, however, that the careful choice of a small number of numerical belief values are sufficient to model the influence of beliefs on the values and as a determinant on action choice. Similarly to CSSMs, beliefs can be perceived from the self, peer or public perspective.

Beliefs are higher level conscious judgments, and we posit that they are less subjected to the phenomena *psychological adaptation* [59] than the values. For instance values such as politeness or dignity perception will tend to return to their average values over timespans of days. Beliefs, however, evolve more slowly, and they do not have natural trends towards average values. This does not mean, however, that beliefs are not affected by timespans without other actions - for instance, the perception of friendship might diminish in the presence of long spans of time without actions reconfirming this friendship.

We model the agent's beliefs using the Dempster-Shafer theory of evidence[60, 61] in the following way:

- the agent's current beliefs are fully encoded in the mass function - no previous evidences are remembered
- incoming evidence can be weighted by significance
- at every incoming evidence, the belief is updated using the standard Dempster's rule of combination (conjunctive merge).
- the value for the positive belief is used as the indicator of the belief.

Although, in general, the semantics of the Dempster-Shafer model is controversial, the results obtained with this model represent a good match to our intuitive understanding of the scene – which, in fact, is what it is exactly what our objective was. We do not want the real probabilities of the events, rather to simulate the algorithms used by humans to maintain their beliefs.

We will use the following beliefs in the modeling of the checkpoint scenario:

B^{SPR}_{threat} the soldiers belief that the vendor itself represents a threat (this does not include the belief that the congestion created by the vendor's presence can represent a threat).

The perceived threat level starts up at a constant value, dependent on the soldier's training and personal perception. In general, the passing of time and human interactions decrease this belief. This belief affects the soldier's judgment of the security level function of the vendor location.

$B^{V \to x}_{\text{unappr}}$ the vendor's belief that the soldier x is unapproachable, i.e. it will not participate in social behavior. This belief starts at a level dependent of the vendor's personal experience, and is decreased by social interaction. This belief affects the vendor's behavior and judgment of possible outcomes of social actions.

$B^{V \to x}_{\text{friend}}$ the vendor's belief that the soldier x is a friend. Friendly actions (casual conversation, exchange of gifts, requests delivered with high mitigation level, lenience in accepting reactions to commands) increase the friendship belief. Actions which are considered rude (unmitigated commands, refusal of gifts) decrease the belief of friendship. The belief also decreases (albeit more slowly) in the absence of friendship maintenance actions (e.g. casual conversation).

$B^{V \to x}_{\text{pubfrnd}}$ the public's belief in the friendship between soldier x and the friend. This belief echoes the vendor's own beliefs but it is updated more slowly, as information propagates from the vendor.

3.3 Experimental Evaluation of CSSM

The proposed model has been implemented using the YAES [62] environment, and a collection of third party visualization tool and the OpenWonderLand 3D virtual environment.

For experiment, we trace five different scenarios, distinguished by different strategies taken by the soldiers at the checkpoint. Each scenario traces the evolution of beliefs and CSSMs over the course of 14 days. These days also model the existence of external factors beyond the control of the soldiers and population: we assume that a medium (orange) alert happens on day 8 and high (red) alert on day 12. In the model we also include action A11 (overnight), that would shift the peer politeness and dignity back to the normal value. We assume that over the weekend, action A11 happens which justifies the rational that a person's dignity is less affected as an accumulative results of bygone days. But the belief is still affected and it maintains the value over the course of interaction.

1. *Rude checkpoint members.* In this scenario, the soldiers of the checkpoint enact a commanding behavior which, due to the use of unmitigated command language and lack of human interaction is perceived as rude by the vendor. This perception is propagated to the beliefs of the general population. The positive side of this scenario is that the perceived security level remains high. However, the perceived politeness is low, the vendor is offended in his dignity, and the public belief is that the soldier and the vendor are not friends. The vendor is incurring some level of financial losses as it will regularly need to occupy unfavorable locations.

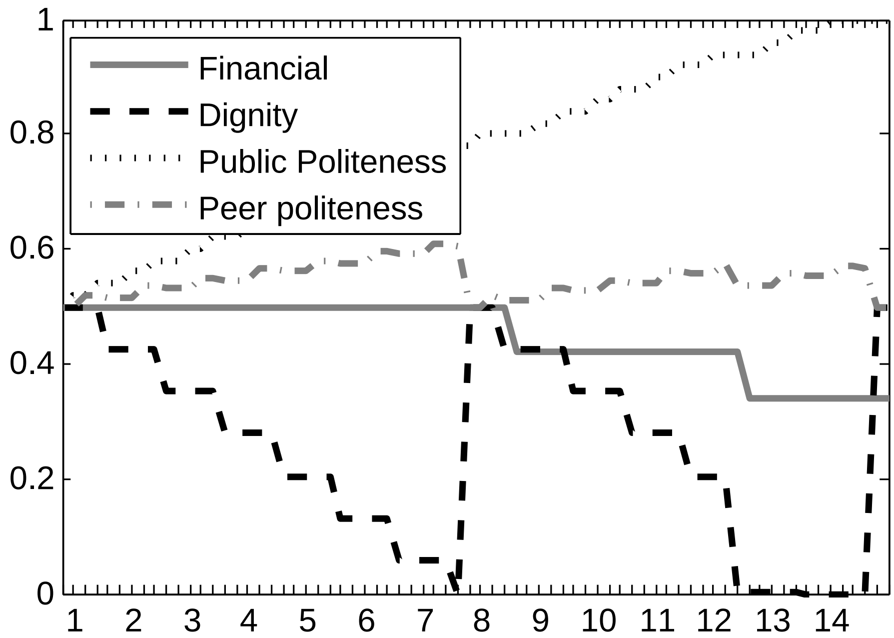


Figure 3.2: Scenario 1 - Dignity (S3) of seller gradually decrease with time due to harsh action (A6, L7) taken by checkpoint members

2. *Overly friendly checkpoint members.* This scenario in contrast with the first scenario has entirely opposite model, representing too friendly behavior of the checkpoint members.

   For instance, when performing action A6 (requiring the vendor to move to a more distant location), the soldiers use highly mitigated speech. At this mitigation level, the seller is free to ignore the command and never moves his cart (even on the high alert days). The scenario is financially advantageous to the seller, maintains a public perception that the vendor and the soldiers are friends. It leads, however to a low level of perceived security.

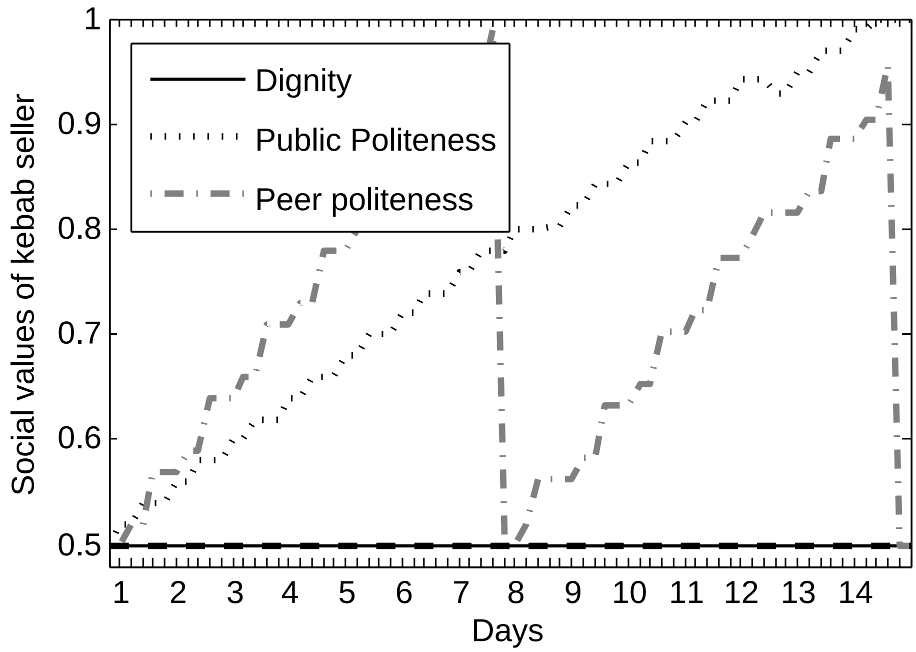


Figure 3.3: Scenario 2 - The peer politeness (S5) rapidly evolves due to extremely polite actions (A6, L1) of checkpoint members

3. *Abrupt changes in behavior.* In this scenario the members of the checkpoint alternative between maximally friendly behavior on days without alerts, with a highly commanding behavior on days with orange and red alerts. One of the unexpected results of this scenario is that the overall friendliness perception is very low, despite the fact that the soldiers are friendly on most days. The reason for this phenomena is due to the fact that a sudden shift to commanding behavior with persons one had established friendship is more damaging to dignity than commanding behavior to a stranger. This scenario, with its abrupt behavior changes, maintain a high level of perceived security, but it maintains a negative overall perception of friendliness.

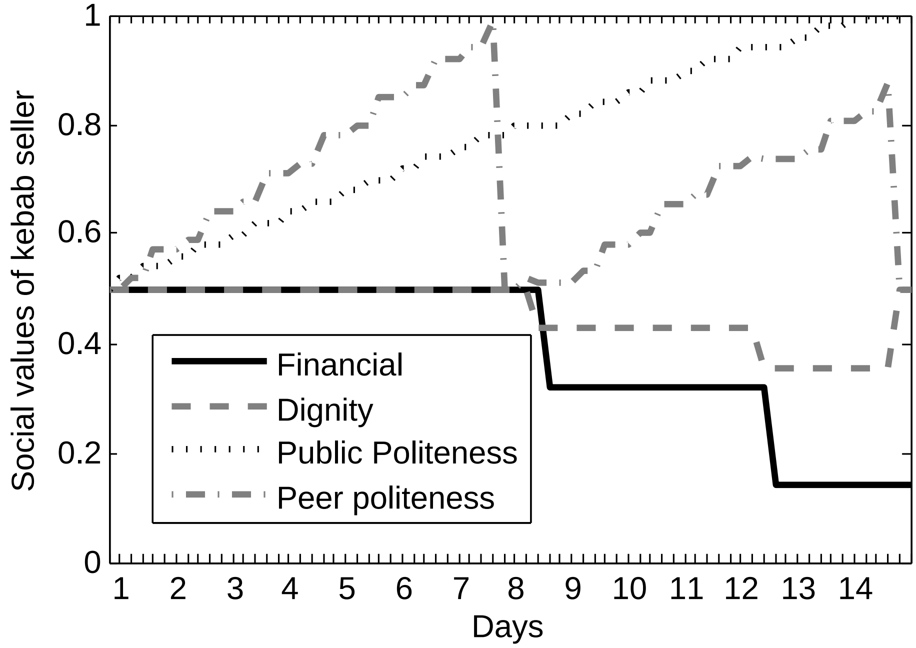


Figure 3.4: Scenario 3 - Two negative bulges can be observed in dignity (S3) of kebab seller on high alert days (day 8, 12).

4. *Moderate changes in behavior.* Similarly to the previous scenario, the soldiers are friendly on days without alarm, while more firm on days with orange and red alarms. This scenario, however, presents less abrupt changes, decreasing the mitigation level of command A6 only until the command is obeyed. They don't have abrupt changes in their behavior and gradually persuade the seller to move over (e.g. increasing directness of speech).

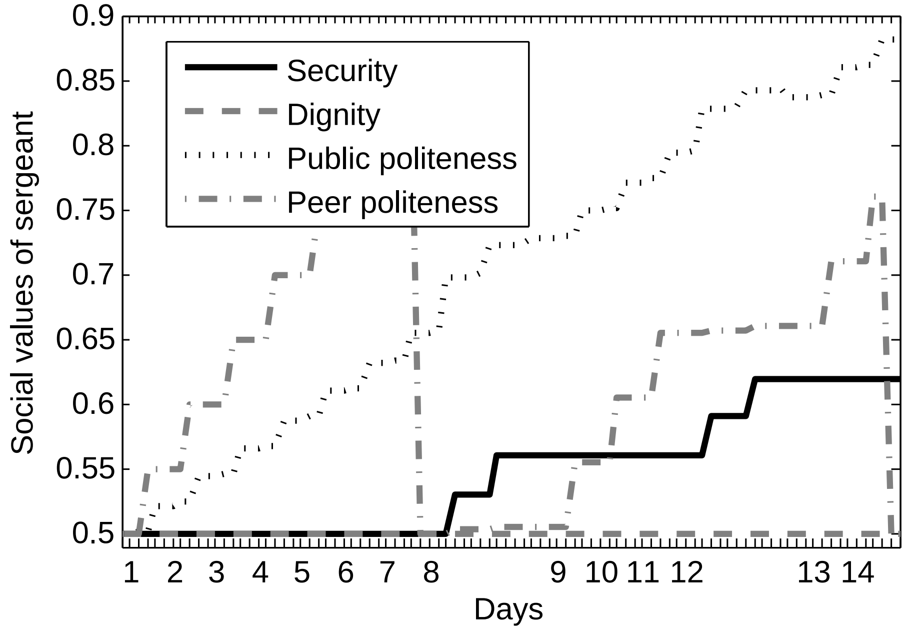


Figure 3.5: Scenario 4 - The security (S2) risk increases with time as the checkpoint members persuade the seller by varying A6 (L1, L3, L7)

5. *Delegation of unpleasant tasks.* In this scenario, we observe the social values of the participants wherein the checkpoint members are assisted by a robot, The sergeant and private are friendly and use low level of mitigation and accept gifts on all the days. On the high alert days, initially they communicate without assistance of robot, but if the seller doesn't move then they send the robot over to perform action A10.

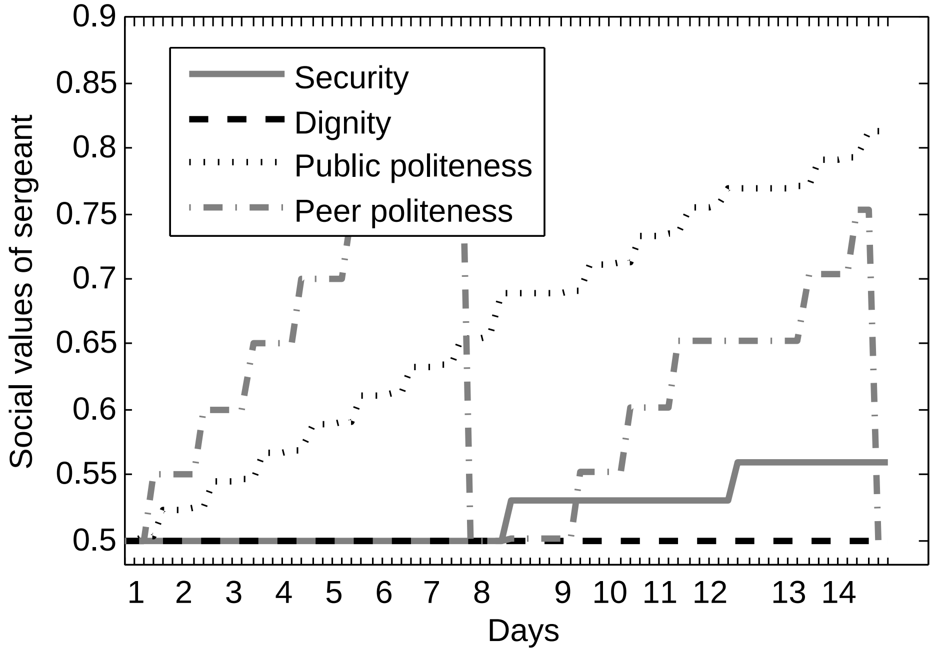


Figure 3.6: Scenario 5 - The security (S2) risk increases as robot assists due to action A2 of seller.

We study the dynamics of the evolution of vendors belief by comparing the belief $B^{V\to x}_{\text{friend}}$ over different scenarios. From Figure 3.7, we can observe the most negative evolution of $B^{V\to x}_{\text{friend}}$ for Scenario 1 (which had mitigated level of speech as L6), and eventually $B^{V\to x}_{\text{friend}}$ drops to minimum level. Further, as contrary to scenario 1, the scenario 2 had absolute positive nature of checkpoint members and the belief of friendship eventually reaches the peak. In scenario 3, a positive trend is observed in $B^{V\to x}_{\text{friend}}$ in the first week. Starting day 8, where there was an abrupt change in behavior of the checkpoint members and also with the accumulative negative behavior on day 12, $B^{V\to x}_{\text{friend}}$ drops to a significant level.

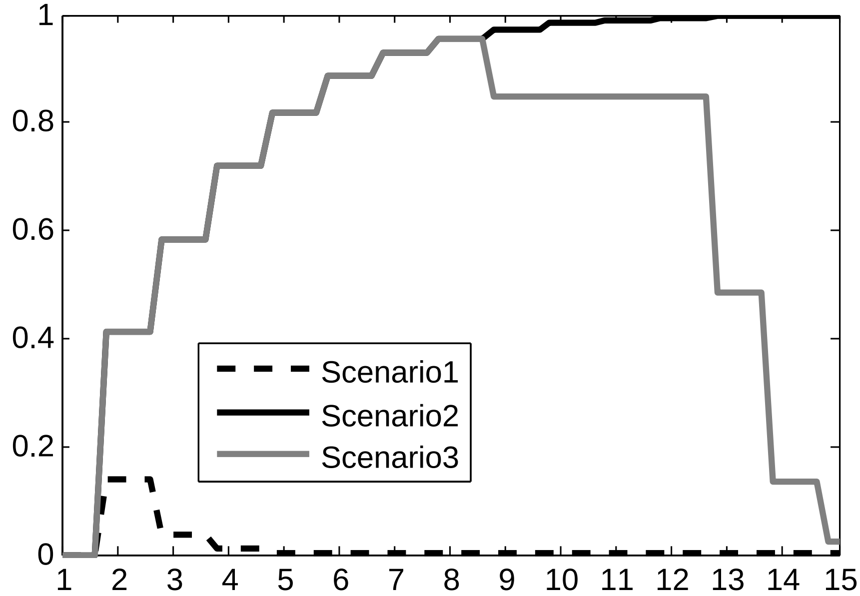


Figure 3.7: The evolution of belief of the vendor $B_{\text{friend}}^{V \to x}$ over different scenarios

CHAPTER 4
SOCIAL CALCULUS - THE USER STUDY VALIDATION

4.1 Modeling the AIFs

An agent acting in a social setting tries to maximize a perceived utility. The contribution of the tangibles and CSSMs to the utility can be complex, non-linear and time-varying. An example of this is the saturation curve provided by the phenomena of psychological adaptation [59]. The change of CSSMs as a result of an action is described by action-impact functions (AIFs). Let us consider a social metric $M_c(A,t)$ showing the value of the metric at time t for agent A. The action-impact function will give the value of the same metric after an action had been performed $a(A_A, A_T, x_1, x_2, \ldots, x_n)$ where A_A is the actor of the action, A_T is the target of the action, and x_i are the parameters which describe the nature of the execution of the action:

$$M_c(A, t+1) = F(M_c(A,t), a(x_1, x_2, \ldots x_n)) \tag{4.1}$$

We need expressions for this function for various agents: the actor, the target, but also their peers. We shall also consider virtual agents which represent, for instance, the public opinion of the bystanders.

The reader may note that our analysis is essentially just a rewriting of the traditional way in which an agent can be built. What is new here, is the *CSSM bottleneck* - we assume that the behavior of the agents in a social-cultural context can be fully described by the CSSMs and the tangible values. The utility function can be, of course, a complex and possibly non-linear function of these values, but it does not depend on anything else. What makes our model more useful for social-cultural modeling is that the components of the utility function are clearly mapped to values which make sense in a certain culture. Finding the AIFs can be seen as a *symbolic regression:* a process through which measured data is fitted with a suitable mathematical formula. Symbolic regression can be performed through manual knowledge engineering. However, there are also several techniques to automatize it, genetic programming being one of the several possibilities.

Genetic programming [63] is an evolutionary algorithm where the individual units of evolution are programs. When applied to symbolic regression, these programs will simply be expressions of the functions we are searching for. GP follows the generic workflow of evolutionary programming. It starts by initializing a diversified population, where each individual unit is represented by a chromosome. For each step, it generates a new set of individuals through the genetic operators of crossover and mutation. Finally, the fitness of the individuals are evaluated, and a selection process takes place, where individuals with higher fitness have higher chances of survival. In the case of GP, chromosomes encode a program, usually in the form of a tree structure. The fitness of a specific program is evaluated by actually running the program over several test cases with known desired outputs.

Poli et al. [64] lists a series of circumstances where GP has been found to show good results. Out of these, there are two criteria which strongly applies to the search for AIFs:

- **The interrelationship among the relevant variables is unknown or poorly understood:** this is clearly the case of the various parameters of human interaction. As we have said above, there is no guarantee that CSSMs form an independent set of variables. In fact, there is normally a strong correlation between the self, peer and public CSSMs.

- **Conventional mathematical analysis does not, or cannot, provide analytic solutions:** there is no mathematical theory behind social calculus. What the assumptions behind the CSSM model say is only that different members of the same culture will evaluate the values similarly. We can make only very loose assumptions about the mathematical form of the AIFs - for instance we can infer that they are monotonic in certain variables, or that they are not periodic in certain variables.

- **Significant amounts of test data are available in computer readable form.** In our case, we have a relatively large data set acquired through our survey. Furthermore, the CSSM assumption that any person immersed in a given culture will provide the same evaluation allows for relatively efficient ways to collect data.

Based on these considerations, we conclude that GP is a good choice for the acquisition of AIFs through symbolic regression.

4.2 Calibration of AIFs

Assigning numbers to social values is an inherently inexact science. However, the working assumption is that the culture enforces a more or less uniform method to calculate the sanctioned social values. This means that we can validate (and, if necessary calibrate) the CSSM model by performing a survey in which persons cognizant with the respective culture will judge the impact on the social values.

In this section, we describe our experience in administering a survey to 91 respondents from various regions in Pakistan. The respondents were presented with several possible unfoldings of the Market Checkpoint scenario and were asked about their personal evaluation of CSSMs at certain points.

The datapoints obtained through this survey will be used as an input into the learning process of the AIFs. Our objective will be that the genetic programming model will evolve functions closely matching those used by the target population when updating their CSSMs.

In the following, we first discuss the problem of the representativeness of the survey, then briefly present the survey methodology and results.

4.2.1 Representativeness of the Survey

One of the important considerations is the representativeness of the survey: are the results of the survey representative of the CSSMs of the target population? It is well known that many academic surveys suffer from the problem of using respondents who are in many ways divergent from the general population and are, in certain ways, "weird" [65].

In the following we will discuss some of the obstacles we perceive in the representativeness of our results.

- The culture of the survey takers (Pakistan) might not be an exact match of the target culture. This is an unavoidable bias - for a perfect localization, one would need to use respondents from the exact geographical location we model.
- There might be a possible misunderstanding between the culture-sanctioned metrics covered by the specific names. Our modeling target was a hypothetical, Arabic speaking Middle-Eastern environment. Our respondents have been primarily Urdu speaking, with a good knowledge of English, and many with at least some level of Arabic. We are confident that the use of English names, together with the Urdu and Arabic translations have provided a sufficiently clear definitions of the values considered (see Table 4.2.2 for some of translations used).
- The distorting factor of social class: the survey subjects have been drawn from a significantly higher social strata (students, engineers, doctors) than the average composition of the market. It is to be determined whether the social class affects the calculations of CSSMs. Our conjecture is that it has only a minimal effect, through secondary implications, which we will outline below.

- The impact of persons cognizant of multiple cultures. Many of the respondents have received some level of Western or Western-style education. It is to be determined whether this impacts their evaluation of the CSSMs. Our conjecture is that is at most a minimal impact. We assumed that people cognizant of multiple cultures are able to evaluate separate CSSMs according to multiple cultures (naturally, within the limit of the cognitive load they can handle). Then, they decide which CSSM-dependent rules of conduct apply in the current situation (which might be a combination of rules), and plan their actions in function of (not necessarily in obeisance to) these rules. This behavior model implies that even people who do not follow rules according to these CSSM settings, will still be able to calculate them.

4.2.2 The Survey Results

The methodology of the survey was as follows:

- the participants were presented with the scenario in a story-board style, with screen-shots and explanation of the ongoing action.

- the participants scored the value of the perceived social value from the point of view of the seller (answering of questions of the type: rate the perceived politeness of the X on a scale of 0 to 10).

The participants were 91 persons from various regions in Pakistan.

While space limits us from analyzing the full output of the survey here, Figure 4.1 shows a representative case. The figure shows the histogram of answers for the public and peer politeness values for action A(7, 5) - order to move using mitigation level 7 and moderate voice level and A(1, 5) using maximally mitigated speech. The graph shows that there is a remarkable consistency in the estimated CSSM values, but also some level of distribution around mean values.

Table 4.1: Names of CSSMs in English, Urdu and Arabic colloquial terminologies

Social Values	Urdu	Arabic
Politeness	مہذب ، شائستہ	التهذيب والسلوك
Dignity	وقار، عزت	احترام الذات
Friendship	دوستی	صداقة
Security	محفوظ	أمن

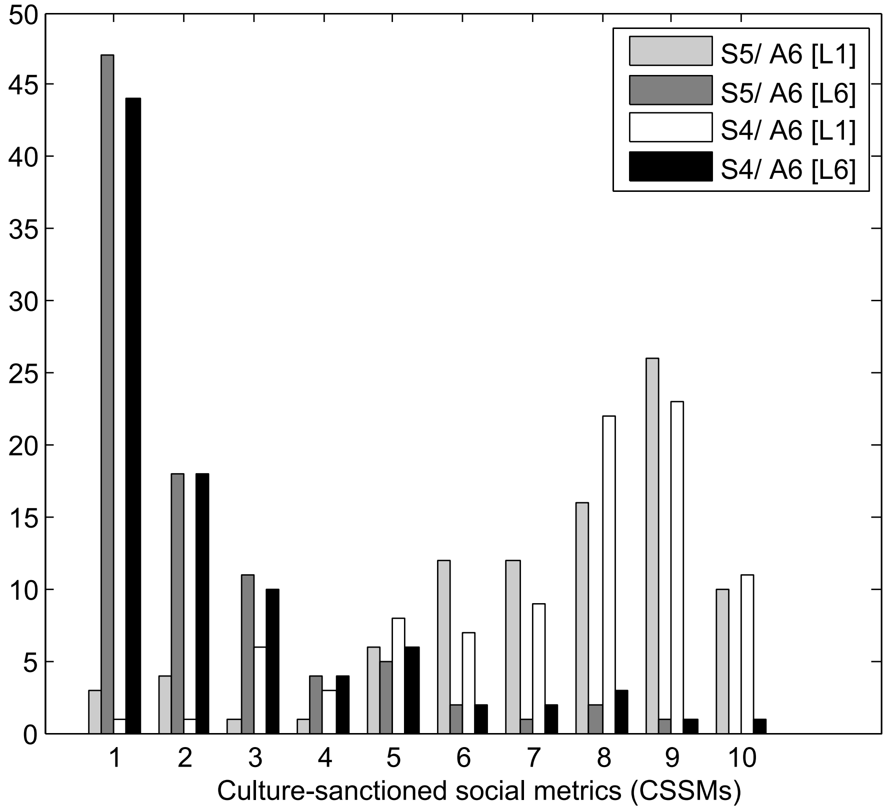


Figure 4.1: The survey histogram for public politeness [S4] and peer politeness [S5] in view of the vendor when the sergeant performs action [A6] (order to move)

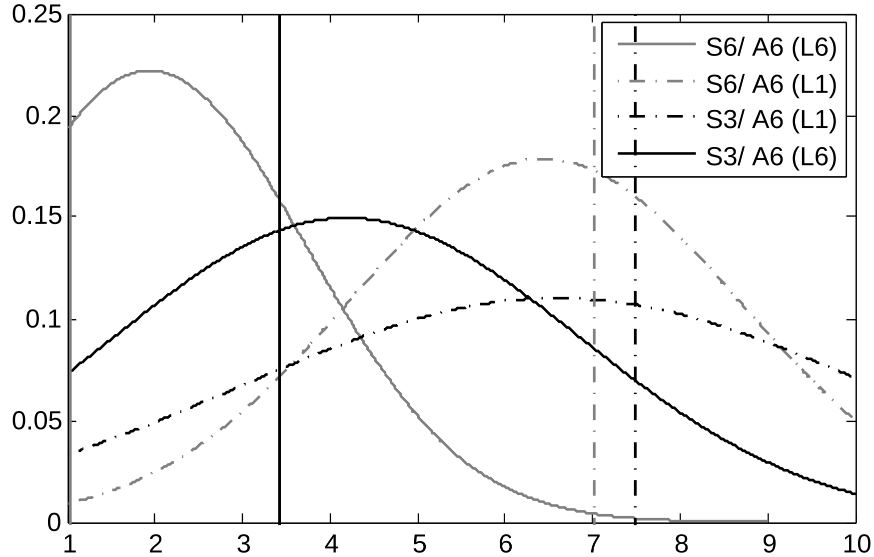


Figure 4.2: Normal distribution for dignity (S3) and friendship (S8) due to actions of order–to-move (A6) with mitigated level of speech L1 and L6

4.3 Symbolic Regression for AIFs

In the following, we will describe the workflow of evolving the AIFs using genetic programming. We will need to specify the function representation (which also defines the structure of the chromosome), the fitness function and its evaluation method, and the genetic operators to be used.

Function representation: To start a GP evolution, we need to define the functional space over which the evolution will take place. In our previous experiments with manual knowledge engineering of the AIFs, we have found it useful to restrict them to a combination of constants, polynomials and Heaviside step functions connected through arithmetic operations. In addition, many GP algorithms use periodic functions such as sine and cosine due to their favorable mathematical properties. This set, however, creates a too wide set of combinations, making evolution difficult. We can use some of our a priori knowledge about the problem domain to make simplifying assumptions about the format of the AIFs:

- the functions are not periodical, thus there is no logical need to use periodical functions such as sine functions
- there is a natural aspect of human behavior which achieves saturation
- with appropriate parametrization, sigmoid functions can both emulate linear functions and the Heaviside step function.

The function set was chosen to include only multiplication, division, addition, subtraction, and a general form sigmoid function **sigmoid (ax - b)**, where all of the three inputs to the function would be evolved using genetic programming.

Besides using the input parameters for the terminal set, we used a set of scalar values having fractional range from 0.1 to 0.9 and decimal range from 1 to 10. The reason for using such values was to evolve the function as a sum of different sigmoid's which helps in regression to evolve better results. For Table 4.3, we can see that the best solution set was evolved using all the combination of all above mentioned terminals.

Fitness function: The evolved functions have been validated by comparing their values with the reference points provided by the survey results. One of the challenges we have encountered was that the survey had been designed to test the CSM assumptions of cultural consistency, not for AIF elicitation. Thus, despite the comparatively large number of respondents, it covered a relatively small set of the AIFs parameter range. To extend this coverage, we have used a cubic spline interpolated surface of the survey results. The fitness function had been defined based on the euclidean distance between the generated AIF and this surface. For future surveys, explicitly designed for AIF elicitation, this interpolation step might not be necessary.

Genetic operators: Finally, genetic operator probabilities, population size, and the number of generations to evolve were chosen through a combination of computational resources as seen from Table 4.3. For our initial phase we used variable genetic operators of crossover and mutation on the population. The values which seemed to guarantee an exploration of the space and diversity in the population while at the same time insuring selection pressure were using low crossover probability and high mutation rate. The formation of new population in this phase was based on the technique that the children would replace the parent population completely, i.e., this option was chosen for using the non-elitist approach (even if children are worse individuals than their parents). From Table 4.3, we can see that the best results were generated using the *tournament selection* for generation of new populations. When the change is relatively small then keeping high level to mutations gives better results in genetic algorithms [66]. For successfully preserving while improving on the solution structure we used low crossover probability and high mutation rate.

One of the problems frequently encountered in genetic programming is *bloat:* the phenomena that the population is gradually taken over by individuals of high complexity (and associated long chromosomes) which offer, at best, minor improvements in fitness. Bloated solutions frequently generalize poorly, and are difficult to interpret by humans. To limit bloat, we have limited the trees to a maximum size of 10. The trees were initially limited to a size of 2 but were allowed to grow only if there was an increase in fitness function.

Table 4.2: Crossover and mutation probability variation using tournament selection for survival

CrossOver probability	Mutation probability	Fitness	Test fitness
0.05	0.95	82.21	24.73
0.1	0.9	58.07	22.2
0.2	0.8	61.1	22.13
0.25	0.75	67.34	33.28
0.5	0.5	65.04	32.5
0.7	0.3	91	91

Table 4.3: Training phase for evolution of CSSMs through genetic programming

PSize	Sampling	Terminals					
		{p}		{p, 10}		{p, 0.1···0.9,1···10}	
		Fitness	Test-fitness	Fitness	Test-fitness	Fitness	Test-fitness
50	Tournament	374.01	120.276	72.02	30.75	273	273
	Roulette	253.1	69.91	91	91	273	273
	Lexictour	379	122.7	74.7	31.47	273	119.3
	Doubletour	380.44	122.9	91	35.1	219	97.9
100	Tournament	350.76	142.81	70.24	31.61	272	119.3
	Roulette	263.17	80.65	90.94	44.76	91	35.1
	Lexictour	260.87	260.87	91	91	219	219
	Doubletour	378.12	163.87	91	91	91	35.1
500	Tournament			51.18	18.75	46.52	14.39
	Roulette	224.60	69.76	74.6955	31.46	219	97.9
	Lexictour	149.03	44.047	50.78	17.71	55.41	22.51
	Doubletour			74.7	31.47	73.97	30.29

Table 4.4: Parameters for the genetic learning

Parameter	Value
Number of generations	75
Population size	50,100, 500
Crossover probability	0.95
Mutation probability	0.05
Function set	{+, -, /, *, sigmoid}
Terminal set	{CSSM inputs, 0.1 $\cdots$ 0.9, 1 $\cdots$ 10}
Selection	{Roulette, Tournament Doubletour, Lexictour}

4.4 Results for the Modeled AIF's

The workflow described in the previous section had been implemented using the GPLab an open-source toolbox for Matlab [67].

In the following we will describe the experimental results for the evolution of the AIF for the dignity CSSM at the action A6 (see Section 3.2.3. This functions has two parameters, the loudness X1 and the offensiveness X2 (the latter being calibrated with the level of mitigation of the speech).

Figure 4.3 shows the evolution of the fitness values of the population during the evolution. Using hard limits on the dynamic size of tree not only helped us in minimizing the bloating effect but also we were able to evolve the functions fairly quickly. Evolving the best equations with the optimal parameters took about 80-120 minutes with a population size of 500 individuals and 75 number of generations. The best fitness was 18.85, using the tournament selection procedure for evolving generations.

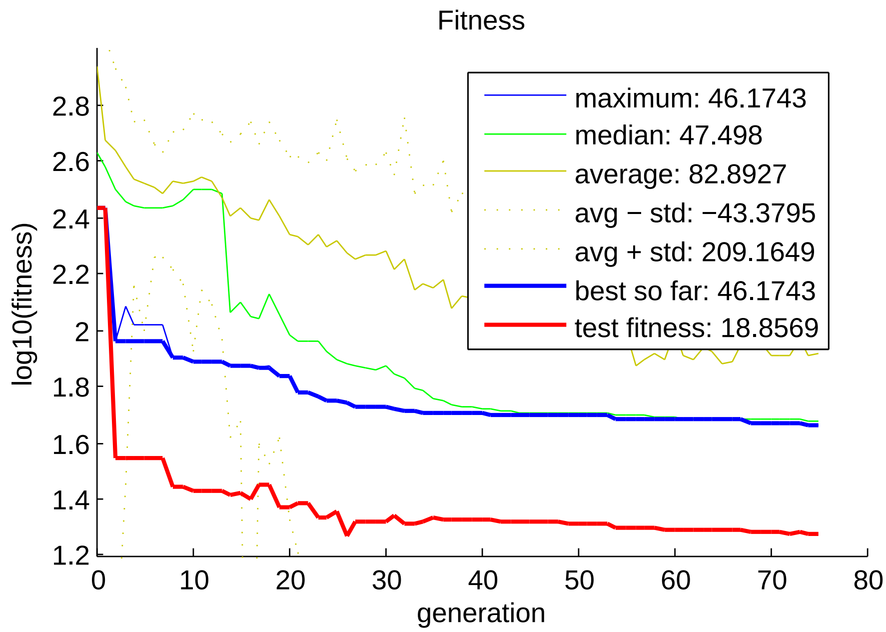


Figure 4.3: The fitness output for best candidate using tournament selection with variable crossover and mutation probability

The best AIF evolved by the system is shown in Figure 4.4.

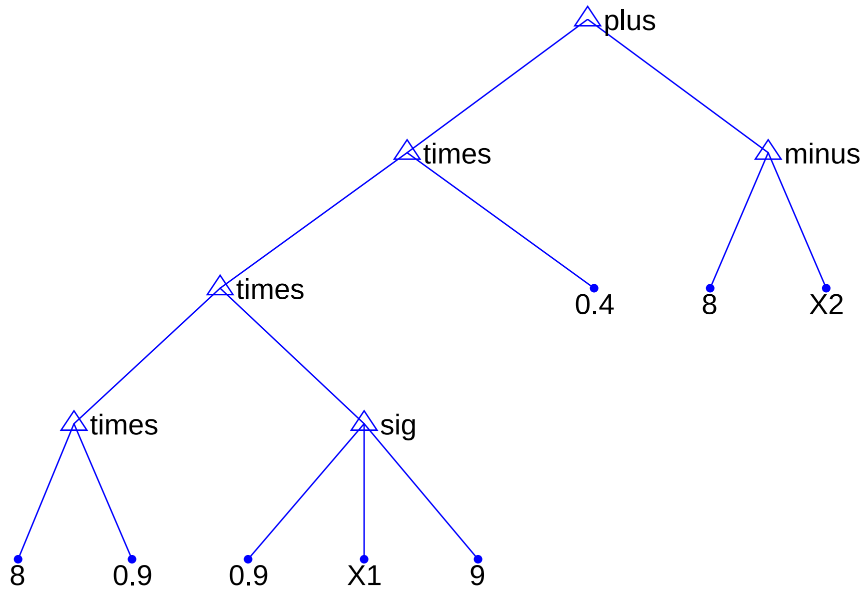


Figure 4.4: The evolved AIF grammar tree for the CSSM dignity in action A6

Finally, Figure 4.5 illustrates the quality of the solution by matching the evolved function to the interpolated data points of the survey. While the match is imperfect (we have survey points both above and below the AIF surface), this appears to be more a result of the inherent noise in the survey data, than the imperfect match. Thus we can conclude that the system had successfully evolved a functional, practically usable form of the AIF for this particular CSSM and action.

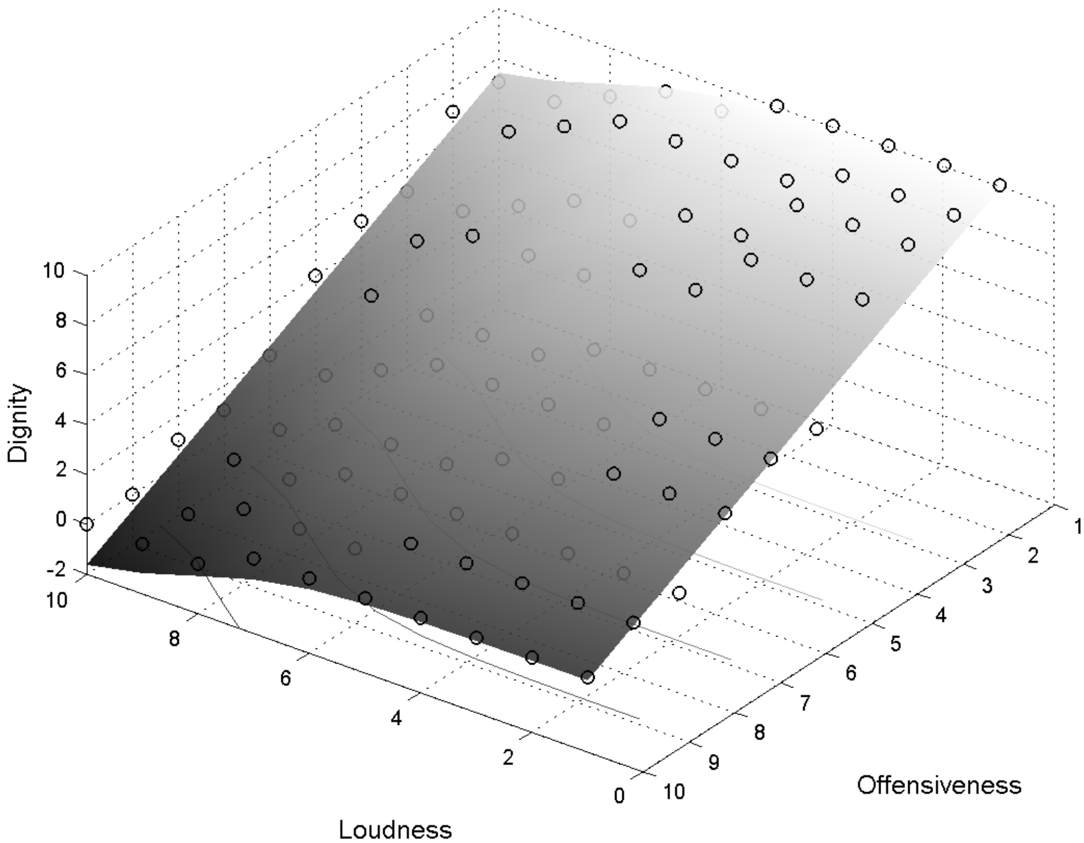


Figure 4.5: A comparison between the evolved function (represented as a semi-transparent surface) and the interpolated survey data (shown as small circles). The small circles *under* the surface are faintly visible due to its transparency.

The evolved output matched with our assumptions about the perceived change in AIF with respect to its variables. The sigmoid (flipped around x-axis) contributes to higher levels of AIF when input variable have low values, which indicates that being polite maintains better dignity. Similarly, we can see that higher levels of x2 (offensiveness) contributes to lower values of dignity.

CHAPTER 5
MICRO-CONFLICT - MOBILE ROBOT IN THE CROWD

We consider the case of an autonomous robot which moves in such an environment. The robot, just like the human participants, has a mission which can be expressed in physical terms. For instance, the mission might be to reach a certain landmark by a certain time, to follow one or more humans at a certain distance, or to maintain its position in a team formation while moving in the crowd. At the same time, the robot needs to stay out of trouble: it must avoid violating the social norms which govern the crowd, taking into account the local social and cultural norms. Part of this can be achieved through path-planning: the robot can plan its path around landmarks and try to avoid dense crowds. Dynamic path replanning, using algorithms such as focussed D*[68] or D*-lite [69] can allow the robot to avoid large, persistent crowds of people. Occasional micro-conflicts with human participants, however are unavoidable, and the urgency of the robots missions makes it unfeasible for the robot to be always the one which "gives way". In general, the robot must avoid violating the social norms, but it should be able to accept some social costs, if it is necessary to achieve its mission.

The scenario considered in our work is as follows. In a busy marketplace a number of customers perform a purposeful movement. They visit various landmarks such as stores and stalls where they spend a certain amount of time, then they move to other landmarks.

We assume that agent movements are independent, *i.e.*, group movement patterns aren't under consideration in this work. The movement of the humans between the landmarks follow planned trajectories, which avoid obstacles, but try to get from one landmark to the other in the shortest amount of time. Reaching their destination in the planned time is the *mission* of the individual human. Delays represent *mission costs*, which the agent tries to minimize. At the same time, the humans need to obey social norms, which require them not to bump into other humans, violate their personal space or block their movement. If they violate these norms, humans incur *social costs*. If two humans are about to collide with each other, they need to take actions to avoid this by one or both of them changing their speed and or trajectory. We call such an encounter a micro-conflict. The strategies of the two agents in a micro-conflict must balance mission costs and social costs. We use the term "micro" to illustrate the fact that such conflicts are normally resolved very quickly (in matter of seconds).

5.1 The Crowd Model for the Marketplace

A busy marketplace might appear chaotic to an outside observer. However, people are not Brownian particles - in fact, each person in the market has a "mission", which we will equate with the task of reaching a goal location L_g which can be a shop, an exit or a location next to another person. Let us assume that an unhindered person aims to reach the goal at time t_g. If the interaction with the crowd creates delays, the actual time will be $t'_g > t_g$. We say that the person incurs a *mission cost* $mc(t_g, t'_g, I_g)$ which depends on the planned arrival t_g, actual arrival t'_g and the importance of the goal I_g. The simplest mission cost formulation is a linear function of $t'_g - t_g$, but other ones, such as deadline dependent formulas are also possible.

The main source of delays when moving in a crowd is the necessity for a person to slow down, stop or alter his trajectory function of the movement of other persons. There is a culture and environment dependent *social cost* sc associated with certain behaviors. For instance, a person incurs a social cost if (a) bumps into another person, (b) violates a person's personal space [70] or (c) blocks a person's movement. For a given person P at a given moment t, we can create a *social cost surface* $sc_P(x, y)$ which associates to every location (x, y) the cost of the person moving there. This surface can be created as a weighted sum of geometrical shapes corresponding to the physical contact zones of the persons in the crowd, their personal distance (1-1.5 ft), social distance (3-4 ft) and their predicted movement cones.

The social and mission costs incurred by a person depend on both his own behavior and those of other crowd members. If everybody would give way to the person, the mission and social costs will be zero. If the person would give way to everybody, the social cost will be zero but the mission cost will be likely significant. In fact, in dense crowds, a person might not make any progress at all if he gives way to everybody. Robots are known to freeze up in dense crowds [33].

Whenever two persons get into a sufficient proximity that social costs are possible in the next step, they need to adjust their movement, each taking into consideration the possible social and mission costs. We call this situation a *micro-conflict*. A micro-conflict is *resolved* when the persons get sufficiently far away that no social costs are possible. We model micro-conflicts with a sequence of one or more 2x2 games where a C move means that the person gives way while D means that it moves forward on his planned trajectory. This model can account for a slow-down (by alternating C and D moves), but it does not cover the options of accelerating or changing the movement path. The payoffs of the game are given by the total costs incurred by the players for the various combinations of moves. The games are not, in general, symmetric, as the cost functions differ from person to person. The games also do not fall into a specific, well known class, as they are dynamically created from the cost surfaces, and each game depends on the outcome of the previous game as well. If two persons are heading on a collision course, they will at some moment encounter some variation of a Hawk-Dove game, where in the case of a (C,D) or (D,C) play the player moving D will have an advantage, but a (D,D) move will have a large cost for both players. A (C,C) move means that neither player moved – the players made no progress, but have incurred mission costs. Thus, even for a (C,C) move, the next game will have different payoffs. It is not necessary, however, for each of the games encountered during the resolution of a micro-conflict to be Hawk-Dove games.

Before moving on to the behavior of a robot, let us first consider how humans "play" the sequence of games in a micro-conflict. Restricting our considerations to a single game,

game theory would tell us to choose a move which maximizes our payoffs with the assumption that the opponent also plays the perfect strategy. As the games in the micro-conflicts are not (in general) zero-sum, this would correspond to a *maximin strategy* (risk minimization). However, this is not an accurate model of human behavior. Crowd participants encounter many micro-conflicts over time, each micro-conflict consisting of several games. Human psychology rewards perceived consistency and predictability. Sudden behavior changes, even if justified on a game theoretic basis, carry their own psychological and social costs. Thus, instead of choosing a strategy on a game-by-game basis, humans choose long term *meta-strategies* which are often associated with their social status. Furthermore, people advertise the type of games they are likely to play by social signals such as clothing, posture and facial expression. Relying on these signs, players can, to a certain degree, predict the moves of the opponents.

5.2 The Micro-conflict

The individual members of the crowd move in a purposeful way: move from one shop to another, stop at various landmarks or head towards the exit along a pre-planned but not rigidly fixed trajectory. We will say that the individuals have a *mission* with a specific value and urgency. The movement of people in such environments is governed by social norms: they are not supposed to violate each other's personal space, block each other's intended direction of movement or physically bump or push each other. The social norms for physical movement depend on the culture and social setting. Different cultures define the personal space of an individual differently, and put different penalty on physical contact. Whether movement in a certain environment can be performed without violating any social norm depends on the density of the crowd: beyond a certain density, an individual which tries to avoid any violation of personal space will not make any advance at all. Groups of individuals moving in dense crowds will enter into *micro-conflicts* if following their planned trajectory would create an unplanned, large social cost through physical collision or severe violation of personal space. The attribute *micro* illustrates the fact that these conflicts are normally resolved in several seconds: one or more participants will alter their speed and/or path, reducing the social cost to an acceptable level.

5.2.1 The Social and the Mission Costs

One way to quantify the decision making process of humans in social settings is by taking into consideration the costs and benefits of certain actions. We will split the cost of movement into the *social costs* depend on the social norms governing the environment and the participants while *mission costs* depend on the specific goals of the human or robot

We will model the social costs of moving in the crowd by a number of geometrical zones associated with the opponent agents. An agent incurs costs whenever it enters into one of these zones. The zones are not necessarily circular, they move and change orientation with the agents. The costs associated with these zones are justified by psychological models of human perception, and they must be calibrated for the individuals as well as for the culture.

The *social costs* depend on the degree an agent violates the physical zones of the agent. In general, a person can avoid occurring social costs by avoiding to enter the specified zones, which normally means giving way to the opponent in micro-conflicts.

For modeling the social costs, we consider three zones:

Physical contact zone: represented by the actual physical size and shape of the human or robot agent. Violating this zone means physical contact and carries a large social cost.

Personal space: is the spatial region which a person (and by extension, a robot) regards as psychologically his [70]. Within the personal space, we model the personal distance (1-1.5 ft) and the social distance (3-4 ft). The cost decreases towards the outside of the area, becoming zero outside the social distance perimeter.

Movement cone: the movement cone represents the space where the human or the robot made public its intention to move. We consider the movement cone as circular pie extending from the agent in the current direction of movement, for a radius equal of 3 seconds movement with the current speed. The movement cone is only relevant for a mobile agent. By violating the movement cone, the opponent forces the agent to change its movement, unless it accepts a high social cost by violating the personal space or even the physical space.

The *mission cost* is proportional to the degree the mission of the person is jeopardized by the actions. We assume that the mission cost is proportional to the delay occurred in micro-conflicts. In some scenarios, for instance, if the person or agent's mission is to reach a landmark before a deadline or to keep up with a moving companion, the mission cost might increase non-linearly with the delay.

We are using a model where the social costs are additive across the cost types and for the multiple agents. For instance, if the agent violates more than one agent's personal space, it will occur the sum of the costs. On the other hand we retain only the maximum social cost for each micro-conflict.

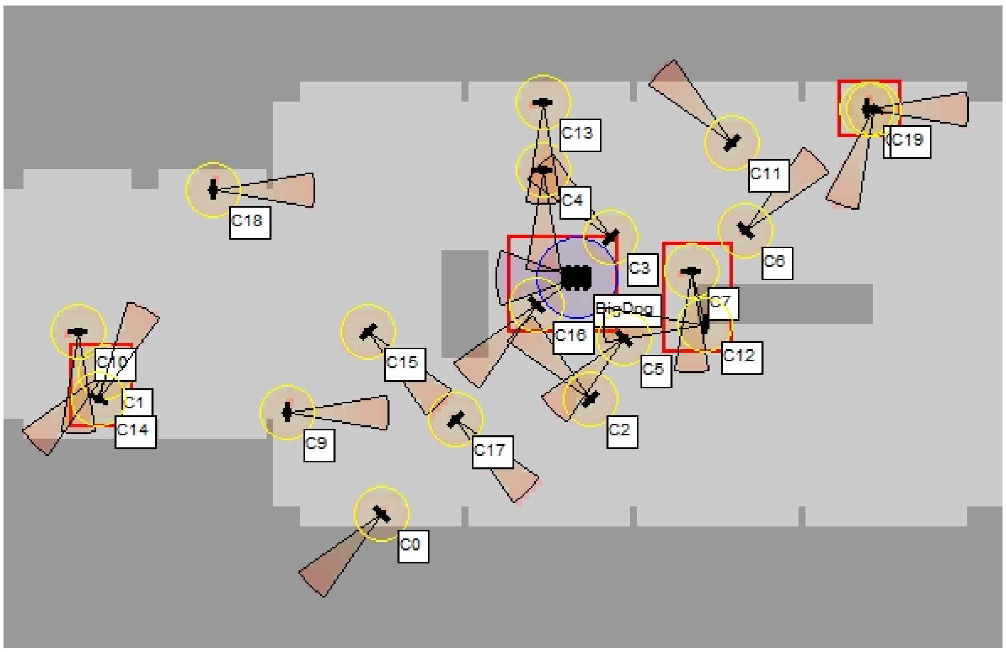


Figure 5.1: A moment in the scenario of the robot navigating a crowd of people on the market. The screenshot shows the visualization of the scenario in the simulator at time $t = 21sec$.

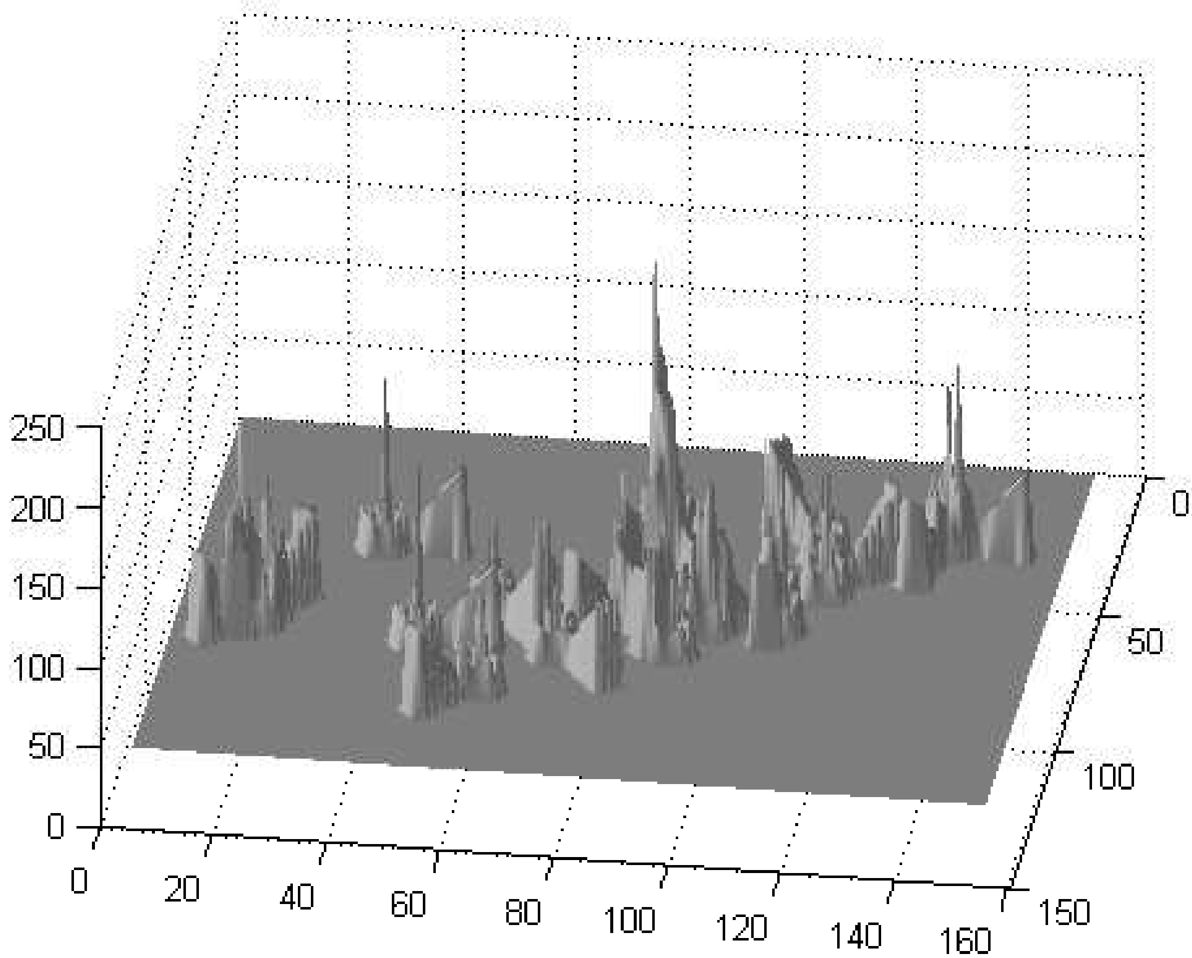


Figure 5.2: The diagram shows the cummulative social cost at that particular moment. The goal of the robot can be interpreted as an attempt to move while keeping to the "valleys" of this constantly changing surface.

5.2.2 Modeling the Micro-conflict

A micro-conflict is a situation in the movement of an agent in the crowd where the next planned action of the agent has a significant, unexpected social cost by violating the zones of one or more opponent. For the current work we will only consider micro-conflicts with exactly two participants. Furthermore, we assume that micro-conflicts will be attended to in the reverse order of their maximal costs (which means in dense crowds, agents will ignore lower stake micro-conflicts until the ones with higher stakes are resolved).

The answer of the agent to a micro-conflict involves the consideration of other alternatives to the currently planned movement: the agent might stop, continue moving with a different speed (faster or slower) or it can replan its trajectory. We model this choice with a two-player one-move game. The move C (collaborate) corresponds to the player stopping, while the move D (defect) corresponds to the agent moving on its currently planned path. This model can account for a slow-down (by alternating C and D moves), but it does not cover the options of accelerating or changing the movement path.

The payoffs of the game are given by the total costs incurred by the players for the various combinations of moves. The games are not, in general, symmetric, as the cost functions differ from agent to agent.

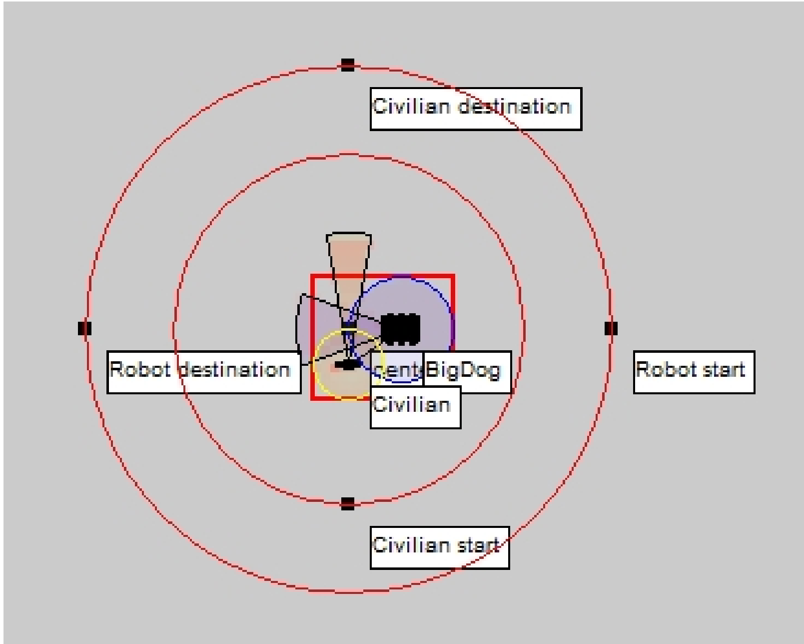
As a note, for these games it is more convenient to speak in terms of cost minimization rather than payoff maximization. Rigorously, the payoffs are the costs with a negative sign.

5.2.3 The Life cycle of a Micro-conflict

A game can be technically created among any pair of agents. However, if the agents are sufficiently far away from each other, the moves (D,D) will have no cost in the game. The agents enter into a micro-conflict when the (D,D) move pair has a non-zero cost for at least one of the agents. The conflict is *resolved* when the (D,D) move pair will have again a zero cost.

A micro-conflict is not necessarily resolved in a single game. It normally requires a series of games, each with a specific set of costs. Even if the two agents play (C,C) which means that they start the next game from the same physical position, the costs of the new game might change if one of the agents has an urgent mission, which would change the mission component of the cost. Figure 5.3 shows the evolution of the games played during a hand-crafted micro-conflict where a robot and a human are heading to a collision course on right-angle trajectories.

It is impossible to predict the nature of the games which will occur during a micro-conflict. The agents heading on a collision course will at some moment encounter some variation of a Hawk-Dove game, where in the case of a (C,D) or (D,C) play the player moving D will have an advantage, but a (D,D) move will have a large cost for both players. It is not necessary, however, for each of the games encountered during the resolution of a micro-conflict to be Hawk-Dove games.



Game at t=8.0 (moves: R: C, H: D)

Civilian BigDog	C	D
C	38.268 1.000	13.500 22.535
D	12.500 1.000	37.268 18.750

Game at t=9.0 (moves: R: D, H: C)

Civilian BigDog	C	D
C	25.768 20.750	2.000 37.268
D	0.000 24.535	24.768 12.500

Game at t=10.0 (moves: R: D, H: D)

Civilian BigDog	C	D
C	25.768 2.000	2.000 0.000
D	0.000 2.000	24.768 0.000

Game at t=11.0 (moves: R: D, H: D)

Civilian BigDog	C	D
C	1.000 2.000	2.000 0.000
D	0.000 2.000	0.000 0.000

Figure 5.3: A hand-crafted single-conflict scenario between a robot and a human. The screenshot of the scenario (above) at time t=7.0 and four individual games at times t=8.0 to t=11.00 as they appear during the resolution of the micro-conflict.

CHAPTER 6
HUMAN BEHAVIOR ADAPTATION IN MICRO-CONFLICT

6.1 A Consistent Strategy for Micro-Conflict Resolution

Mobile robots moving in a crowd need to conform to the same social standards as the human participants. Imitating human behavior is a natural choice in these situations - however, not every human behaves in the same way. On the other hand, it is known that humans tend to behave in a consistent way, with their behavior predictable by their social status.

We are tempted to think that making a robot behave in a socially acceptable way is equivalent for the robot to mimic "human behavior". However, if we observe human social settings, we find that not all humans behave in the same way in all social encounters. First, human social behavior has a certain randomness even for seemingly identical settings. Second, humans vary their behavior in function of the opponent and the circumstances of the encounter. And finally, not every human choose to obey the social rules. On the other hand, it is a well known fact of psychology that the overall functioning of the social life depends on the *consistency* of behavior. One of the principal requirements of human social interaction is that the participants form a theory of mind of each other [71]. This allows them to predict the beliefs, goals and actions of the interaction partner. This allows for a significant variance on allowed behavior. However, a certain consistency in the behavior is required, as we cannot model or predict the mind of an erratically behaving interaction partner. The agents need to take this into consideration about their human interaction partners; furthermore they need to act such that the humans can form a predictive model of them. Research show that humans are willing to treat agents as social actors [72] although in some situations they will treat humans differently from agents or robots [73]. In most social settings, it is not the individuals pursuing aggressive or defensive strategies who are causing the most social disturbance but the ones who are erratically switching between the two.

In this chapter we develop strategies for the resolution of the micro-conflict games introduced in the previous chapter. The micro-conflicts are modeled as a series of two-player games, in which the participants must deploy specific strategies. The consistency of the behavior does not mean that every human deploys the exact same strategy every time (in fact, such an overly uniform strategy creates problems in which symmetrical strategies can be broken only by one party abandoning the game). Rather, a consistent behavior means that the behavior in the micro-conflict can be predicted from observable attributes of the participant.

6.2 Human Opponent Strategy

6.2.1 Strategy consistency and choice of strategies

We define *strategy* as the algorithm used by an agent to determine its choice of move in a given game. Restricting our considerations to a single game, game theory would tell us to choose a move which maximizes our payoffs with the assumption that the opponent also plays the perfect strategy. As the games in the micro-conflicts are not zero-sum, this would correspond to a *maximin strategy* (risk minimization).

However, this is not an accurate model of human behavior, because the human game-playing strategy takes into consideration other factors beyond the current game. First of all, crowd participants will encounter many micro-conflicts over time, each micro-conflict consisting of several games. Human psychology rewards perceived consistency and predictability and there is a social cost of being perceived in having an erratic behavior.

Second, beyond the micro-conflict games costs, the agent's behavior must be consistent with other social values such as dignity, politeness, "face" and other metrics. The first implication of all this is that instead of choosing a strategy for the individual games, the agents will choose *meta-strategies* which they will follow consistently across the games of the micro-conflict. Meta-strategies can contain stochastic elements and considerations of factors outside the current game (such as the history of the games in the micro-conflict or predictions of future games). The existence of stable meta-strategies means that the players can, to a certain degree, predict the moves of the opponents. Under these conditions, maximizing is not an optimal strategy - stochastic expectation maximization strategies can yield a better value in the long run.

The next questions involves whether humans use mixed strategies in micro-conflicts. It is well known that for Hawk-Dove games the only symmetric Nash equilibrium is a mixed strategy equilibrium. On the other hand it had been argued that the randomness involved by mixed-strategies is not the normal way for humans to operate: humans do not perform mental coin-tosses, and even if they would want to, they have difficulty generating random outcomes without external physical means. In the particular case of micro-conflicts, however, we can safely assume the existence of mixed strategies as there is sufficient randomness both in the lack of knowledge about the exact game (the Harsányi interpretation [74]) as well as in the uncertainty about the strategy of the opponent [75].

6.2.2 Modeling human meta-strategies

One of the characteristics of human meta-strategies is that humans enter into micro-conflicts with a clear view of what type of resolution they would prefer. These strategies not only determine the behavior of a specific human player, but they also provide information to the other players. The use of a specific meta-strategy in the case of a human is signalled through the physical movement itself. In human-to-human interaction there are a number of other means through which this communication can happen: there is a priori information that can be inferred from social status, previous acquaintance and physical characteristics. In addition to this, human players can perform communication during the micro-conflicts using social signaling[76] or even natural language. These communication means, however, are not available for human-to-robot interaction.

We will consider four meta-strategies. For each, we will describe the *intent* of the agent A when encountering agent B, followed by its expression in terms of costs. The intent of an agent will be to either cooperate (C) or to defect (D) in a micro-conflict.

MS1 **Respectful:** I am going to give B a wide berth. Agent A tries to avoid any social cost in the interaction with B, playing C for all games unless the predicted costs are very low.

MS2 **Tight-after:** I am going to let B pass, but pass very close behind him. This can be modeled by a stochastic model where the agent plays with a high confidence that the opponent plays D (i.e., the assumption that B will cooperate is 0.25).

MS3 **Tight-front:** I am going to cross in front of B (but will avoid direct physical contact). This can be achieved by a stochastic strategy that weights the opponent's predicted choice with a high confidence that the opponent plays C (in our model, we assume a probability of 0.75).

MS4 **Bully** The agent decides to minimize its mission costs, ignoring almost all social costs. The assumption behind this model is that this behavior will make the opponent play C, thus keeping the costs low.

These meta-strategies can be transformed into a specific mixed strategy for each individual game encountered by the agent in the resolution of the micro-conflict. Note that although these meta-strategies are not optimal, certain combinations can yield near optimal social costs for the overall micro-conflict. The encounter between a bully and a respectful agent will yield a low social cost through the restraint of the respectful agent. However, the series of C moves by the respectful agent implies a high delay and thus a high mission cost for it.

Another observation is that the high level intent in the meta-strategy might not necessarily be accomplished. If both agents use Tight-front, naturally, only one of them can pass first. What will happen is that depending on the geometric configuration, there will come a moment when the other agent's cost for the D move will outweigh all other considerations, and it will need to play C, allowing the other agent to pass first. Nevertheless, the series of moves will be different from that of an agent which would have played Tight-after.

6.3 Mobile Robot Strategy

The intent of an agent is not physically observable unless observations are made from past experience of micro-conflicts. The motivation behind the robot's strategy is to be consistent with its behavior during its interaction with different types of agents in social context. The robot's strategy uses a two-fold approach: there is a passive phase and an active phase. In the passive phase the robot performs the offline learning of a classifier which helps to decide the intent of the agent. The robot is trained with samples of micro-conflicts performed by humans. In the modeled system, some of the physical features of the human agents partially overlap. Hence, this fuzziness in the physical attributes of humans helps in introducing noise to the robot's identification system. In the active phase the mobile robot, with the help of trained classifier selects an appropriate micro-conflict strategy.

There is a strong motivation for the robot to play a meta-strategy which is, at least superficially, similar to that of humans. Furthermore, if the robot can make the assumption that the human will play a particular consistent meta-strategy chosen from a limited set (such as the MS1 . . . MS4 strategies outlined above), it can try to infer what strategy the opponent uses and choose an advantageous counter-strategy. As we have seen, human players have various means of social signaling to communicate their chosen strategy. If the robot lacks the ability to communicate in a similar way, it needs to rely exclusively on the information gleamed from game-play.

We have implemented a Naïve Bayes classifier that allows the robot to adapt its behavior to the opponent, by probabilistically predicting the next move of the opponent, and using it to weight the costs of its own moves. For instance, if the agent classifies its opponent as a Bully, the agent only needs to consider the costs of the (C,D) and (D,D) move pairs, knowing that the opponent always plays D. The Naïve Bayes classifier takes the form

$$p(X = x_i \mid Y = y_j, Z = z_k) \quad = \quad p(X = x_i \mid Y = y_j) \tag{6.1}$$

where x_i is the category of the human crowd member, y_i is the set of physical attributes and z_k is the social context. The mobile robot can be trained on any set of observable features. In our case, we divide the crowd members into six classes as shown in Table 6.1.

Table 6.1: Modeled attributes of human crowd members for the Middle Eastern social context

Physical Features			Micro-conflict Strategy	Classification
Chin-type	Hair-type	Height (ft)		
Round	Short	1 - 4	Tight-front	Child (Male, Female)
Square, Heart	Short, Long	5 - 7	Respectful	Adolescent (Male, Female)
Square, Heart	Bald, Long	4 - 7	Bully	Senior (Male, Female)

Among the other advantages of using an adaptive model is that it also helps the robot to adapt itself according the urgency of the mission. For example, for urgent missions the robot would try its best to minimize time-cost whereas for normal missions it would try to minimize the social-costs. This can easily be integrated in the model, by training the robot to vary its consistent strategy for different opponents. We model the consistent meta strategy using Naïve Bayes defined as follows:

MS5 **Consistent:** The robot, with the help of a classifier would classify the human opponent based on physical attributes such as face and body features. The robot will select an appropriate meta-strategy that will likely be successful based on the classification of the human opponent, and maintain this strategy for all the games until the micro-conflict is resolved.

6.4 Micro-conflict Strategy Evaluation

In the following we describe the results of a series of experiments that test the behavior of the consistent micro-conflict resolution strategy. The experimental setup models a marketplace with a narrow space surrounded with shops whose entrances serve as landmarks, as well as internal obstacles. A number of shoppers perform purposeful movement, which involves visiting shops for a shorter or longer times. The path chosen by the individuals balances the shortness of the path with the avoidance of the obstacles and large groups of people. Micro-conflicts are resolved through a succession of games.

In this baseline scenario, we consider the presence of a patrol of peacekeeping soldiers traversing the market while being accompanied by a Boston Dynamics Big Dog robot [77]. The mission of the robot is to follow the soldiers through the crowd as closely as possible with consistent behavior towards the population. The soldiers can change their movement at any time, triggering frequent path re-plannings, for which we use the D*-lite algorithm [69]. The robot participates in micro-conflicts in the same way as the human participants. Naturally, the robot's personal space and physical space is different from that of a human (a Big Dog robot is larger than a human).

We consider two different sets of experiment for the the behavioral simulation of the robot. In the first set of experiments we evaluate different set of populations against the consistent strategy of the robot. In the second set of experiments, we evaluated different strategies of robot against a single population set of the humans. In both scenarios, the humans agents play consistent sets of strategies against each other based on their social status.

6.4.1 Varying the Population Metrics

We consider three distinct times of the day (morning, afternoon and evening). We assume that the population of the marketplace varies in function of the time of the day - each part of the dayis dominated by a particular age group. For each part of the day, we consider dense male population: the population is uniformly distributed with 70% males.

The population-set (PS) of agent in the *morning* has the majority of the agents from the senior age group. The population-set for the *afternoon* is dominated by the agents of children age group and in the *evening* the population is dominated by the adolescent age group. The distribution statistics for each time of the day is as follows

PS1 **Morning** 10% Children, 20% Adolescent

PS2 **Afternoon:** - 30% Senior, 30% Adolescent

PS3 **Evening:** - 70% of Adolescent, 10% Children

6.4.2 Varying the Micro-conflict Strategies for the Robot

For different meta-strategies, to compare the intervals of the incurred social cost and the mission cost, we run the each experiment twenty times. In each experiment, we vary the meta-strategies [MS1 $\cdots$ MS5] for the robot in the experiment. The comparative analysis of various meta-strategies helps to determine the effectiveness of a consistent meta-strategy in different scenarios.

6.4.3 Training the Naïve Bayes Classifier

For the consistent meta-strategy [MS5], the robot uses a Naïve Bayes classifier which is trained with a set of six hundred examples using the data-set generated from Table 6.1. The height attributes of training set used for Naïve Bayes training set has the following statistics:

Table 6.2: Height attribute from the training-set for the human classes

Attribute	Class-Type					
	Child		Young		Senior	
	M	F	M	F	M	F
Height (mean)	2.49	2.48	6.06	5.96	5.06	5.01
Height (std. dev.)	1.063	1.0722	0.7851	0.8237	0.8224	0.8426

After classifying the human opponent, the robot selects the counter strategy in a micro-conflict using the social graph schema (Figure 6.1).

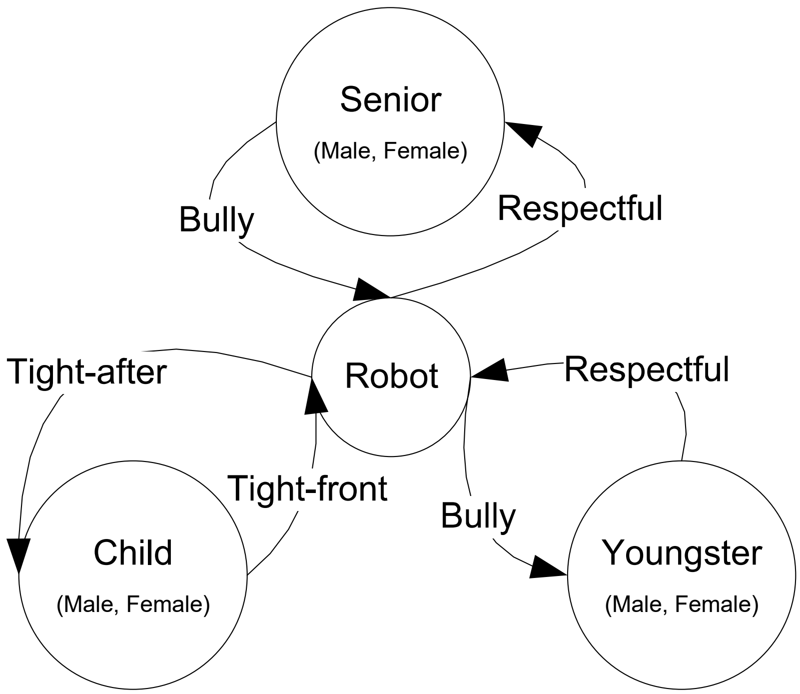


Figure 6.1: Social Behavior Graph for the robot

6.4.4 The Experimental Results

6.4.4.1 Morning

For the first experiment, we consider the morning when there are more seniors in the market-place. Following the social graph schema (Figure 6.1), the robot is expected to be considerate towards the seniors during micro-conflicts and hence it should incur low social costs during the *morning*. From the incurred social costs results (Figure 6.2), we observe that indeed the robot cooperates more during the morning in the marketplace.

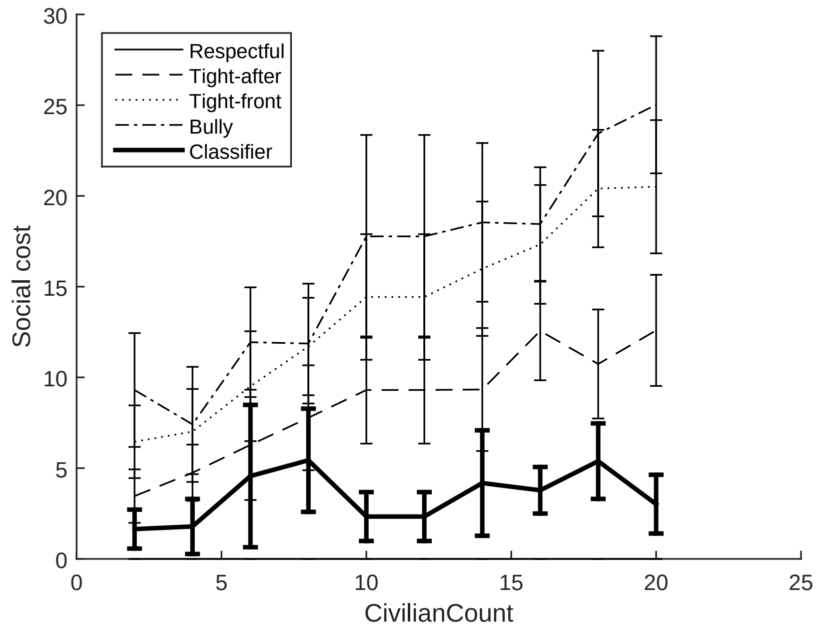


Figure 6.2: Social cost incurred during the morning time

The results for the mission cost incurred during the morning is shown in Figure 6.3. We can observe that besides being respectful during most of its micro-conflict, the robot will employ the *bully strategy* and *tight-after* strategy whenever possible. Hence, the mission cost is lower as compared to the meta-strategy where robots cooperates during all of the games. Another observation is the increase in mission cost with the increase in the density of the crowd (the robot cooperates more and incurs more mission cost). Hence, the robot is adapting its behavior with the crowd variance during the morning session.

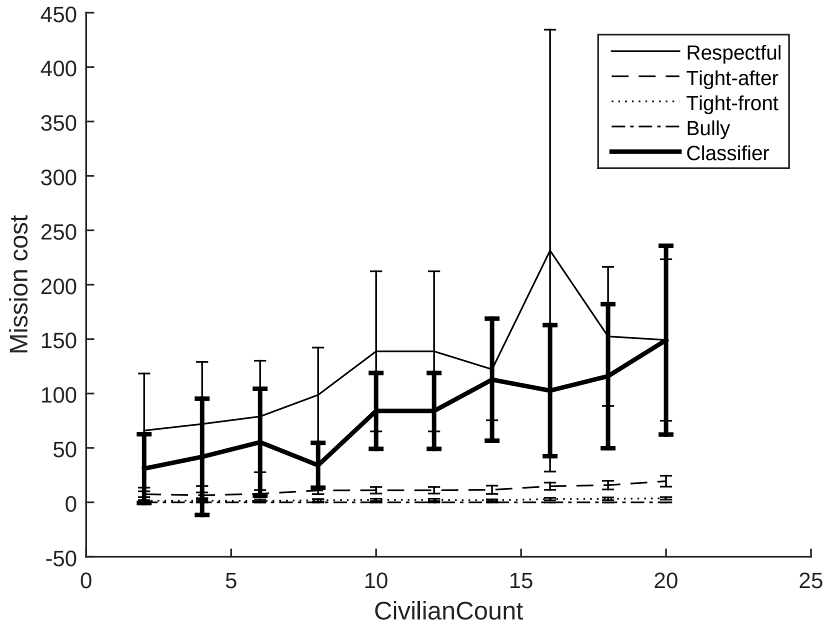


Figure 6.3: Mission cost incurred during the morning time

6.4.4.2 Afternoon

During the afternoon, the marketplace has an equal mix of adolescents and seniors - 30% each. There are more children in the marketplace during the afternoon as compared to the morning and the robot expects that the children would cut the robot's path from the front during the micro-conflict, i.e., children would use meta-strategy [MS3] *tight-front*. Comparing the social cost incurred during the afternoon (Figure 6.4) to the social cost incurred during the morning (Figure 6.2), we observe that with the increasing number of civilians, the social costs increases during the afternoon. The reason is the increase in the number of adolescents and children and the mobile robot uses strategy of *tight-after* (in micro-conflicts with children) and *bully* (in micro-conflicts for adolescents).

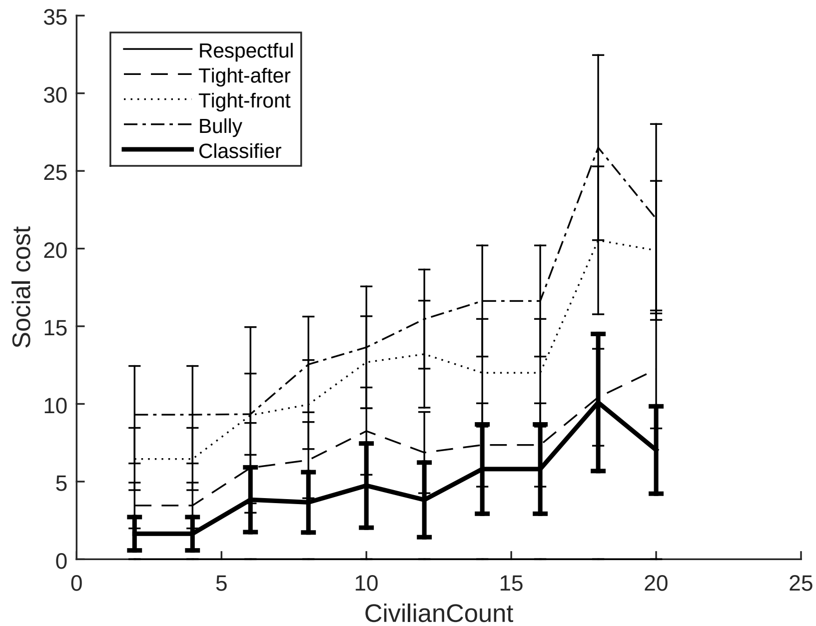


Figure 6.4: Social cost incurred during the afternoon

Comparing the mission costs incurred in the morning (Figure 6.3) to the mission costs incurred during the afternoon task, we observe that with the increasing number of civilians the mission cost remains consistent. The rational for incurring a low but consistent mission cost is due to more bullish behavior of the mobile robot during its micro-conflicts.

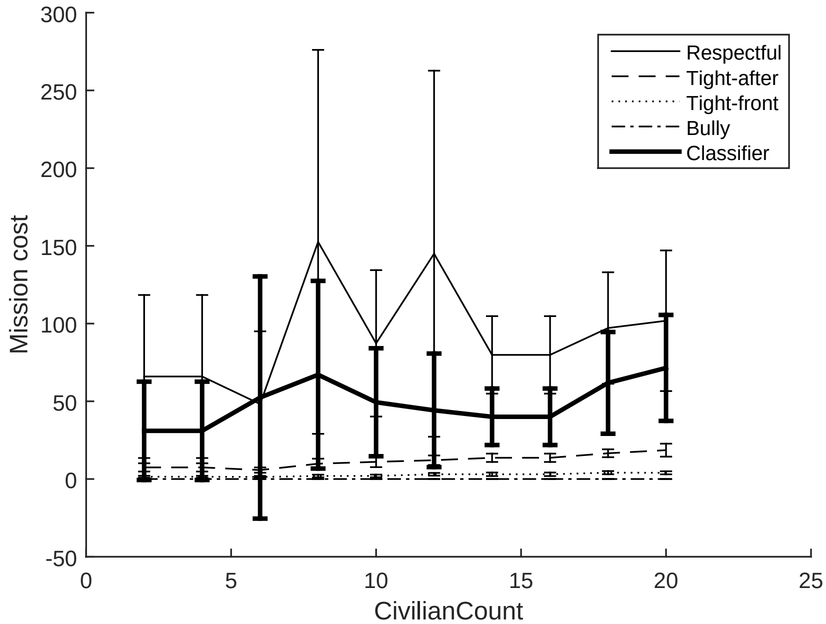


Figure 6.5: Mission cost incurred during the afternoon

6.4.4.3 Evening

Figure 6.6, shows the results of the social cost incurred during the evening. We can observe that the incurred social cost is more in the evening as compared to the morning but the incurred social cost remains consistent with the increase in the number of civilians. The reason is the fully cooperative behavior of the adolescents, which unlike the children do not come in close proximity of the robot (i.e. adolescent cooperate and do not use meta-strategies of defect or tight-front).

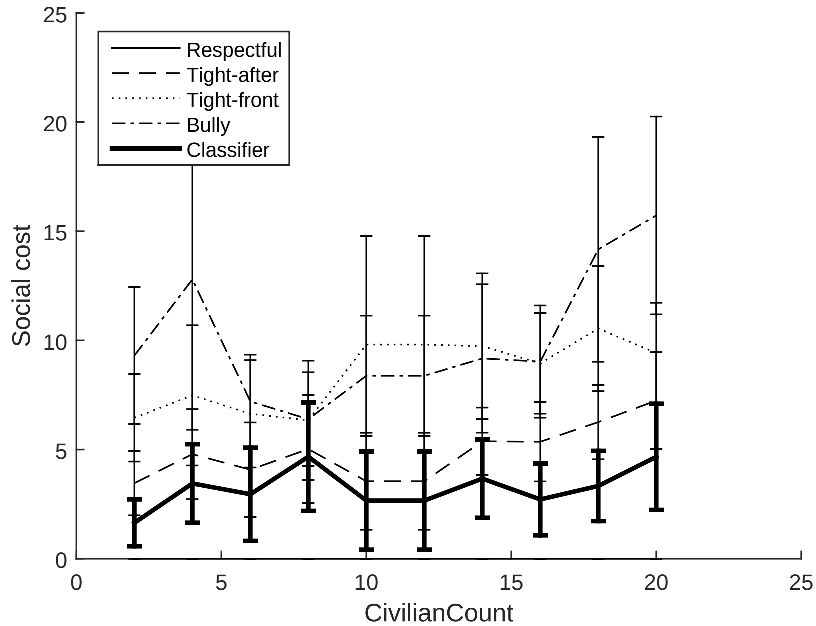


Figure 6.6: Social cost incurred during the evening

The mission cost incurred in the evening is the least compared to other times of the day. In the evening the robot expects the adolescents to cooperate and hence the mobile robot will defect in most of the micro-conflicts. Therefore, as the mobile robot does not cooperate often in the evening, the incurred mission cost for the mobile robot will decrease respectively.

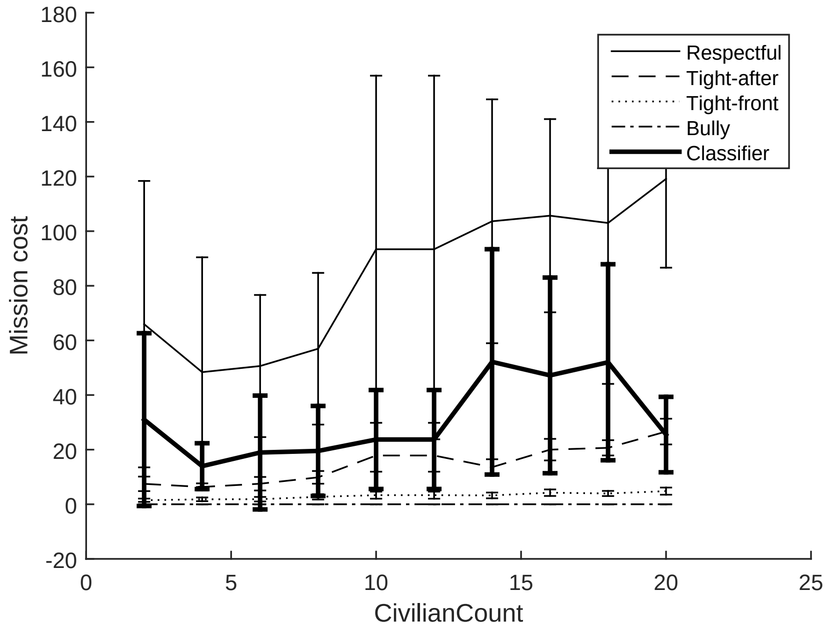


Figure 6.7: Mission cost incurred during the evening

CHAPTER 7
HUMAN BEHAVIOR IMITATION IN MICRO-CONFLICT

7.1 Learning Human Behavior in Crowds

Moving in a crowd requires a balance of assertiveness and politeness. While it is impolite to invade other people's personal space or cut off their movement, in dense crowds forward movement is impossible without a credible threat of personal space violation. Our objective is to develop an autonomous behavior for a mobile robot which imitates the decisions performed by a human controller. We model the situation where participants in a crowd must decide who has the right of way as a micro-conflict resolved through a sequence of games where a C move means that the player gives way, while a D move means it continues to move forward. We collect data from human controllers navigating a robot and resolving micro-conflicts in a simulated marketplace. These recordings are then used to learn a micro-conflict resolution strategy which imitates the human controller's behavior. Through a user study, we find that observers can not distinguish between the fully autonomous and the remote controlled robot's behavior.

The ability to navigate a dense crowd of people is an important human skill. The appropriate behavior depends on the culture and specific circumstances: even the highly polite Japanese will behave assertively and without deference when they need to catch a subway train. While we can learn certain principles from guidebooks, ultimately a traveler in a foreign country must observe and adapt to the local customs of in-crowd behavior.

Let us now consider a mobile robot controlled using the principles of adjustable or mixed autonomy [78, 79]. In such systems, there is a human remote operator who exercises nominal control over the robot. However, depending on the circumstances, the behavior of the robot might alternate between different degrees of autonomy: from teleoperation, to waypoint methods, goal biased autonomy and finally, fully autonomous behavior where the robot is able to set its own goals. Even moderate shifts towards autonomy can reduce the cognitive load of the operator or allow a single operator to control multiple robots.

We consider a mixed autonomy robot moving in a dense crowd in a busy marketplace. The robot has an urgent mission and uses the D*-lite algorithm [80] for navigation. However, the robot also has another mission, *the social mission*, which requires it to act in accordance to the local customs of crowd movement. Whenever a robot violates a social norm, for instance, by entering the personal space of a pedestrian, or colliding with him, the robot will incur a *social cost*. It is impossible to avoid all social costs when moving a dense crowd. Instead, the robot needs to make rational decisions to balance the social and mission costs. We model this decision problem as a *micro-conflict* - a game that is played by pedestrians and/or robots whose movement affects the outcome [23, 81, 82].

For our work, we use imitation learning to create an autonomous robot controller which behaves in micro-conflicts similarly to human remote operators. We started by performing a user study to record the decisions made by humans remote operators in micro-conflicts. In each micro-conflict the robot would either (a) stop and allow pedestrian to pass or (b) will continue to move along the same path. The human remote operator had to select the robot's decision in specific situations, while the system recorded the incurred social and mission costs. We used supervised learning algorithms on the recorded data to model three decision-makers: (i) ensemble learning using random forests, (ii) support vector machine (SVM) and (iii) neural network learning using an evolutionary computational model. For the training phase appropriate features were selected by the use of information gain (IG) and principal component analysis (PCA). After reducing the dimensionality of the recorded data, we trained the random forest classifier and SVM classifier with the pruned data. In the case of neuroevolution, NEAT [83] has an inherent property of reducing the dimensionality of the neural network. Hence, for training NEAT, we used all of the features of the recorded data without applying dimensionality reduction. The decision-making abilities of the models were tested for four unique scenarios. We analyzed the abilities of the learned decision-makers by comparing their output to the decisions made by the human participants for the same scenarios.

7.2 Data Collection for Imitation Learning

7.2.1 The User-study for Imitation Learning

The objective of our work is to emulate the behavior of the robot remotely controlled by a human operator. Our starting point is a mixed autonomy system where the robot assumes responsibility for the navigation while the operator is asked to take decision when the robot enters a micro-conflict. We thus need to collect the decisions made by the operator, and features describing the circumstances in which those decisions were made. The collected data is specific to the operator's "style" and to the environment in which it was collected (football crowd, oriental market, Japanese train station).

For the experiments, we created a simulated environment where the operator of a Big Dog robot [77] must navigate a simulation of a crowded market. The shoppers perform purposeful movement and use a mix of consistent meta-strategies to resolve micro-conflicts. We will assume that the costs (negative payoffs) for players are the sum of their respective social and mission costs.

With this framework, we collected a number of "runs", with the goal that the behavior of the operator will serve as a demonstration or imitation target for the fully autonomous robot. The null hypothesis would be that the operator takes decisions exclusively based on the game payoffs. We have seen however, that people in the crowd do not use exclusively the game payoffs as the basis of their decisions, but also consider the consistency of their own strategy, environmental circumstances and so on. In fact, our strategy of learning to imitate the human operator is based on the assumption that the operator does not simply apply an optimal risk minimization strategy (that could be calculated by the robot without any need of learning) and that the strategy applied by the operator is part of a consistent meta-strategy (which can be learned and applied to games not previously seen by the robot).

The user study was conducted with the help of 12 participants. The human player was presented with a visual interface which provided an overview of the environment and a keyboard-based control of the robot. In order to separate the social behavior from navigation skill, the robot used a mixed autonomy control: the navigation of the robot remained under the control of the agent, with the user receiving control only in situations when the robot entered into a micro-conflict with a crowd member. The presence of the micro-conflict, the personal space and the intended movement cone of the participants had been clearly indicated on the screen. When receiving control, the player could select between moving forward (corresponding to playing D) and staying (corresponding to playing C). Albeit the payoffs of the game had been calculated and used by the adversaries to adapt their play, the game matrix had not been presented to the user, who was instructed to play based on visual feedback as if he was driving the robot through remote control and an overhead camera. To avoid incorrect readings due to the limited reaction time of the user, we stopped the simulation until the user selected a choice, after which the simulation resumed at normal speed.

In a typical experiment the user had to drive the robot across the busy market, during which it encountered about 10-12 micro-conflicts, each being resolved with 2-8 games. We created 5 different scenarios with different crowd sizes {20, 30, 40, 50, 100} and runs were recorded as video files (see Figure 7.1). The number of games played depended partly on the human subject, as more assertive players cleared a micro-conflict in a smaller number of games (but potentially, incurring higher social costs).

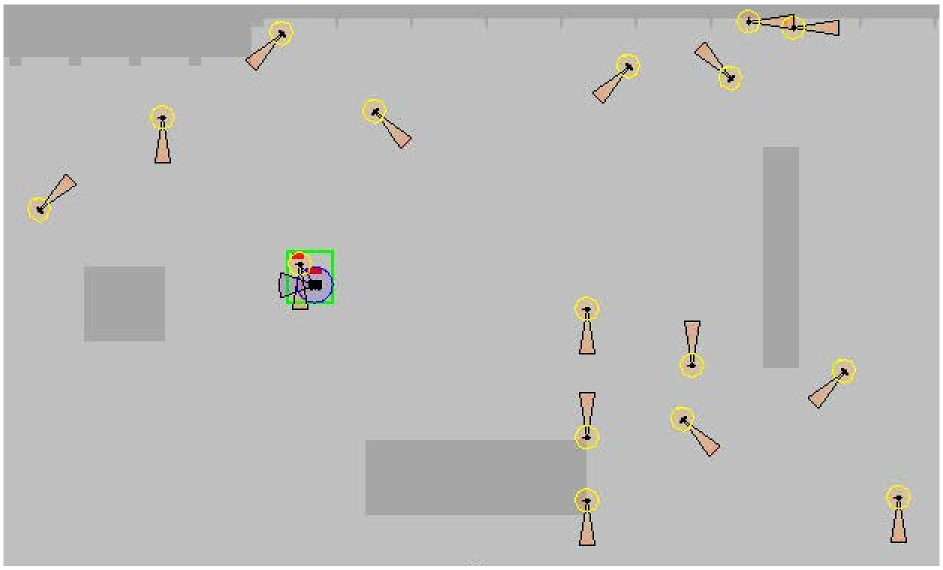


Figure 7.1: Scene from the recorded video with 20 crowd members (sparse crowd). The dark grey boxes are static obstacles. For each crowd participant we show the physical space, personal space and intended movement cone. Active micro-conflicts are indicated by rectangles surrounding the participants.

7.2.2 The Dataset for Imitation Learning

Our working assumption is that the behavior of humans in micro-conflicts can be described as a two-player game, in which the payoffs for both sides are calculated as the sum of the social cost and the mission cost for the player corresponding to a given move (we call this the *composite game*). Nevertheless, the question remains whether this is a good model for the behavior of a human user controlling the robot. For instance, if the robot operator completely ignores the game payoffs, the proposed modeling approach is useless. To investigate the factors used by the human players in their decisions, we decided to collect a larger set of features together with the action C (the robot temporarily stalls its motion) or D (the robot continues its current motion) taken by the user:

7.2.2.1 The robot's own payoffs

R_{CC}, R_{CD}, R_{DC} and R_{DD}. These are the values based on which the user would act if he would be playing an incomplete information game (with no information about the opponent's payoffs).

7.2.2.2 The opponent's payoffs

O_{CC}, O_{CD}, O_{DC} and O_{DD}. The 8 values O_{xy} and R_{xy} would be used by the user if he would be playing a perfect information game.

7.2.2.3 External values

These are values which are not part of the payoffs of the given game. If the user takes these values into consideration, it means that its strategy is not optimal for the given game, but influenced by external factors. Observations of people behaving in crowds validate that such considerations exist: for instance, a person might give way to four passersby but angrily cut in front of the fifth one. The five external values which we hypothesized might possibly affect the behavior of the human agent are listed in Table 7.1. These include the absolute values of the mission and social cost for this particular micro-conflict and the maximum social cost collected by the agent before the current micro-conflict. We have also included the time delay incurred by the agent before the current micro-conflict t_d.

To signify the importance of social cost on the overall decision, we also included the predicted social cost before start of micro-conflict. Further, to maximize the socially enacted information we attributed one of the features with the maximum social cost endured by the robot till that point. Another feature was the *time delay* which helped us in including the temporal impact on the overall decision making. The last feature which we felt was important, keeping in view the urgency of the mission, was to include the robot's mission cost. The list of features is given in Table 7.1.

Table 7.1: The features collected for modeling the imitation decision maker

R_{xy}	The payoffs of the composite game for the robot
O_{xy}	The payoffs of the composite game for the opponent
N_d	The N number of civilians at distance d at the start of micro-conflict
$\delta_{mission}$	The mission cost before start of the micro-conflict
δ_{social}	The social cost for this particular micro-conflict
δ_{smax}	The maximum social cost for the robot before the start of the current micro-conflict
t_d	The time delay cost for the robot before the start of micro-conflict

7.3 The Imitation Learning Framework

We have used three different learning techniques to create controllers imitating the human behavior. These controllers act as classifiers which classify situations into those where the robot should move C versus D. The overall learning process for the three techniques is illustrated in the flowchart in Figure 7.2. The ensemble learning and the SVM based approaches share a significant part of the learning pipeline and will be discussed in this section. The neuroevolution-based learning process has a different pipeline and will be discussed in the next section.

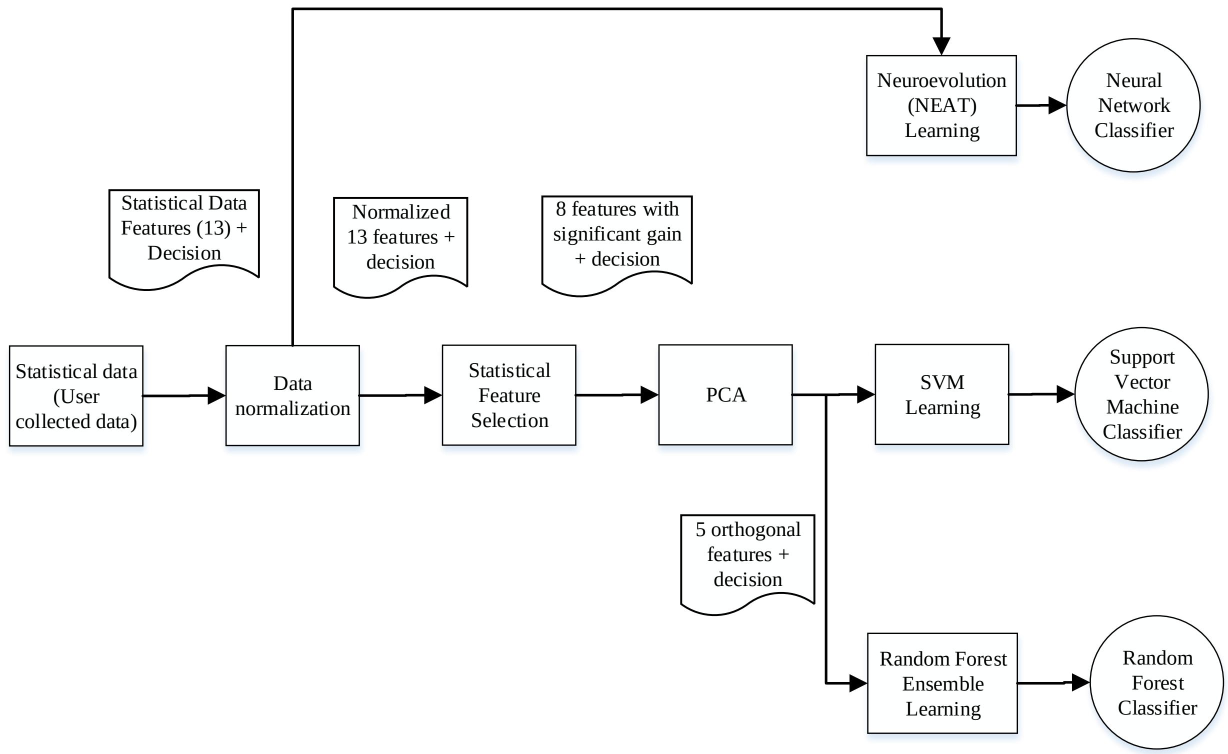


Figure 7.2: The flowchart of the learning process

7.3.1 Imitation Learning using Ensemble Learning and Support Vector Machine

7.3.1.1 Pre-processing the Data

7.3.1.1.1 Data normalization The 13 features collected have their own native, incompatible data ranges expressed in terms of mission cost, social cost, combinations of the two (for the game payoffs), time delay and crowd density. To maximize the efficiency of the learning, these values had been normalized to the [0,1] range, based on the range of the samples in the collected dataset.

7.3.1.1.2 Feature Selection Using Statistical Information As a first step in the selection of the features, we measured the information gain provided by the individual features. Let $\vec{F} = \{f_1 \dots f_n\}$ be the feature vector for n number of features and let $\mathbf{X} = x1 \dots xk$ be the k normalized instances in the training dataset. The information gain for i-th attribute is given in Equation 7.1.

$$IG([C,D], f_i) = H([C,D]) - \{H([C,D])|f_i\} \tag{7.1}$$

where H is the information entropy and f_i is the i-th feature of $\mathbf{X}$.

Calculating the information gain for the features had shown that the most valuable features were t_d (IG=0.249), $\delta_{mission}$ (IG=0.212), O_{DC} (IG=0.208), R_{CC} (IG=0.166) and δ_{social} (IG=0.0567).

The information gain for the remaining features were zero or near zero. The fact that the information gain for N_d was zero, shows that the human subjects did not consider the crowd density when making decisions in micro-conflicts. The fact that R_{CD} and R_{DD} turned out to be zero means that the human subjects did not consider their own costs in the event of the opponent defecting. While the games in micro-conflict are dynamically created, the most challenging games are the ones which have the structure of a hawk-dove game. In these terms, it appears that the human subjects were assuming the opponents to be "doves".

Overall, the information gain analysis step led us to discard five attributes with negligible information gain: R_{DD}, O_{CD}, R_{CD}, δ_{smax} and N_d.

7.3.1.1.3 Dimensionality Reduction Using Principal Component Analysis After the dimensionality reduction step, we still had a number of 8 features, some of them likely correlated. Thus, instead of further pruning the features we opted for dimensionality reduction by using principal component analysis PCA [84]. PCA helps in dimensionality reduction by providing with a set of new attributes that are linear combination of the original attributes. These new attributes have eigenvectors formed using orthogonal transformation, thus they are statistically uncorrelated. The principal components of the data will be the eigenvectors associated with the largest eigenvalues. The higher dimensional data $x_k \in \Re^i$ is projected into lower dimensional vector $y_k \in \Re^j$ (where $j < i$). Hence, given the mean $\mu = \frac{1}{K}\sum_{i=1}^{K} x_i$, the linear projection gives us

$$y_k = C^T(x_k - \mu) \tag{7.2}$$

Here C^T is the transpose of list of eigenvectors which were selected of the basis of highest eigenvalues from the covariance matrix of the preprocessed dataset. Table 7.3 shows the resulting principal components of the data. For our case, as the features were measured on different scales having variance in them, we used opted for the correlation matrix (shown in Table 7.2) for PCA transformation.

Table 7.2: The correlation matrix of the attributes from principal component analysis

	R_{CC}	R_{DC}	O_{CC}	O_{DC}	O_{DD}	δ_{social}	$\delta_{mission}$	t_d
R_{CC}	1	0.35	-0.07	0.04	-0.07	-0.07	-0.08	-0.08
R_{DC}	0.35	1	-0.05	-0.02	-0.05	-0.02	-0.03	-0.02
O_{CC}	-0.07	-0.05	1	0.53	0.97	-0.11	-0.11	-0.11
O_{DC}	0.04	-0.02	0.53	1	0.53	-0.05	0	0.02
O_{DD}	-0.07	-0.05	0.97	0.53	1	-0.08	-0.1	-0.11
δ_{social}	-0.07	-0.02	-0.11	-0.05	-0.08	1	0.79	0.74
$\delta_{mission}$	-0.08	-0.03	-0.11	0	-0.1	0.79	1	0.99
t_d	-0.08	-0.02	-0.11	0.02	-0.11	0.74	0.99	1

Table 7.3: The formulated eigenvectors and there corresponding eigenvalues. Note that each eigenvector contains at least 0.95 variance of the pruned dataset. These components are only computed to provide transformed datasets for the classifier.

Eigenvalues	Eigenvectors
2.81332	$0.53\delta_{mission}$ + $0.522t_d$ + $0.481\delta_{social}$ - $0.301O_{CC}$ - $0.296O_{DD}$ ...
2.28364	$0.542O_{DD}$ + $0.538O_{CC}$ + $0.439O_{DC}$ + $0.273\delta_{mission}$ + $0.269t_d$ ...
1.33126	$-0.697R_{DC}$ - $0.692R_{CC}$ - $0.148O_{DC}$ - $0.061t_d$ - $0.059\delta_{mission}$ ...
0.67271	$-0.661R_{DC}$ + $0.597R_{CC}$ + $0.394O_{DC}$ - $0.155O_{DD}$ - $0.154O_{CC}$ ...
0.56402	$-0.757O_{DC}$ + $0.384R_{CC}$ + $0.303O_{DD}$ + $0.302O_{CC}$ - $0.26R_{DC}$ ...

Therefore, after transformation we obtain vector $\mathbf{Y} = \{y_1 \ldots y_k\}$, where the k-th instance had v_n number of transformed features.

7.3.1.2 Classification

Once the dataset had been reduced in dimensionality, the next step is to develop a classifier that for any given set of features would classify them into situations requiring a C or D answer. Once trained, this classifier can be directly used as a decision-making engine for a robot resolving micro-conflicts. The overall effect is one of imitation learning: what the robot learns is not to act in an optimal way, rather to imitate the decisions of the human subjects who were used to collect the dataset.

The classification problem is essentially a supervised learning problem over the five features obtained after applying feature selection and PCA as described above. We have experimented with a number of supervised learning techniques. The best results had been obtained using random forests, an ensemble learning technique and support vector machines. In the following we describe the application of these techniques to our learning problem.

7.3.1.2.1 Using Ensemble Learning for Classification The basic idea behind decision trees is to use multi-level decision systems that would sequentially classify the instance using features associated at each level until we reach a final decision. Hence, the feature space is separated into distinct regions in a sequential manner.

One of the known problems with decision trees is its number of sibling variants. The reason for variance is linked with low generalization of decision trees for the training data-set used for their construction. If an error occurs high among the nodes it propagates downstream affecting the leaves. One way of improving the generalization error is to use bootstrap aggregating (bagging). The main idea behind bagging is to create a number of M variants $\mathbf{Y_1}, \mathbf{Y_2}, \ldots, \mathbf{Y_M}$ of the original dataset $\mathbf{Y}$. Each set $\mathbf{Y_i}$ is created by uniform sampling with replacement from the dataset $\mathbf{Y}$.

The random forests technique uses bagging to create trees with random feature selection. We have implemented the technique using the Weka [85] library. For an M-tree random forest classification, using features $\{v_1, v_2 \ldots v_n\}$ the classifier takes input y_i and $M - trees$ assigns the label C or D. Being an ensemble learning method, the output is based on the majority vote of the sub-classifiers.

7.3.1.2.2 Using Support Vector Machines (SVM) An SVM classifier searches for hyperplane which separate the classes using a maximum margin to allow for generalization. The hyperplane takes the form:

$$g(\mathbf{y}) = \mathbf{w}^T\mathbf{y} + w_0 = 0 \tag{7.3}$$

where the direction of the hyperplane is decided by $\mathbf{w}$ and the position is determined by w_0. The goal for SVM is to find the direction which can give us maximum margin. Thus, for our case of binary classification, we represent the decision C with 1 and decision D with -1 such that:

$$\mathbf{w}^T\mathbf{y} + w_0 \geq 1 \qquad \forall \mathbf{y} \in C$$

$$\mathbf{w}^T\mathbf{y} + w_0 \leq -1 \qquad \forall \mathbf{y} \in D$$

However, in our case the two classes are not separable classes. For these situations, we can formulate the SVN as a cost minimizing optimization problem [86]:

$$J(\mathbf{w}, w_o, \varepsilon) = \frac{1}{2} \parallel \mathbf{w} \parallel^2 + \delta \sum_{i=1}^{N} \varepsilon_i \tag{7.4}$$

Hence we get:

$$\begin{aligned} \text{minimize} \qquad & J(\mathbf{w}, w_o, \varepsilon) \\ \text{subject to} \qquad & d_i[\mathbf{w}^T\mathbf{y}_i + w_0] \geq 1 - \varepsilon_i \end{aligned}$$

where $d_i \in \{1, -1\}$ and $\varepsilon > 0$ are slack variables. The slack variables are used as the measure of error in the misclassified $\mathbf{y}_i$. If $\varepsilon = 0$, then $\mathbf{y}_i$ is correctly classified. If $1 > \varepsilon > 0$, then $\mathbf{y}_i$ is correctly classified but close to the margin. If $\varepsilon > 0$ then $\mathbf{y}_i$ is misclassified and $\delta \sum_{i=1}^{N} \varepsilon_i$ becomes the penalty term. Here δ is the externally set penalty cost associated with the misclassified vector.

For the two instances $\mathbf{y}_i$ and $\mathbf{y}_j$ we use the radial basis function given as:

$$K(\mathbf{y}_i^t, \mathbf{y}_i) = exp(-\gamma \parallel \mathbf{y}_i^t - \mathbf{y}_i \parallel^2), \gamma \text{ ¿ } 0 \tag{7.5}$$

where $\gamma = 1/2\sigma^2$. Equation 7.5 defines a spherical kernel with center $\mathbf{y}_i^t$ and radius γ. We will use cross-validation to determine the appropriate values penalty cost δ and the kernel radius γ. For the implementation of SVM we used the LibSVM library [87].

7.3.1.3 Cross-validation and Overall Accuracy of the Classifiers

Cross-validation was used to calibrate and test the random forest and the SVM classifier. Using k-fold cross-validation process, we initially divided the training set into k equal bins. We performed k runs and during each run we sequentially trained the model on $k - 1$ bins and tested it for the remaining bin. For experiments, we choose $k = 10$, a commonly recommended approach.

The random forest module was tested using tree range = {10, 15,. . ., 40}. Each of those were constructed while considering 3 random features. We selected the random forest which provided us with the minimum error on the cross validation results. The maximum depth for the trees was set to have no bounds. The out-of-bag error (OOB) for the 35-tree random forest was 0.2743. For comparing the results of the best forest tree we performed 10-fold cross-validation, which is not usually required as one can get a good estimate of the random forest from OOB. The confusion matrix after performing validation on original data-set and then performing validation using 10-fold cross validation is given in Table 7.4.

For the SVM-classifier, we had to find the combination (δ, γ) for the best penalty cost δ and the value of gamma γ. Therefore, we used grid selection [88] which suggests using the exponentially growing sequence of $\delta = \{2^{-5}, 2^{-3}, \ldots 2^{17}\}$ and $\gamma = \{2^{-15}, 2^{-13}, \ldots 2^{5}\}$. We used 10-fold cross-validation on each model trained based on the combination (δ, γ).

From results we observe that as we increase δ from $2^{-5} \rightarrow 2^{15}$, the accuracy increases from 58% to 60% for $\gamma = 2^{-15}$, then it drops for $\delta = 2^{17}$. And if we increase γ from $2^{-15} \rightarrow 2^{3}$, then the accuracy increases from 60% to 75% for $\delta = 2^{15}$, and then it drops for $\gamma = 2^{5}$. Hence, from results we choose, $\gamma = 2^{3}$ and $\delta = 2^{15}$. The confusion matrix for the SVM-classifier is given in Table 7.4 where $\bar{C}$ and $\bar{D}$ are the classified outputs .

Table 7.4: Confusion matrix for controllers evolved using Random Forest and SVM classifier

	Random Forest		*SVM*	
	$\bar{C}$	$\bar{D}$	$\bar{C}$	$\bar{D}$
C	144	32	170	6
D	44	83	70	57

Both random forest and SVM classifier had an output accuracy of approximately 75%. However, SVM appears to have a tendency to misclassify D instances as C (while almost never making the opposite error). In contrast, the random forest module makes both types of errors with roughly the same probability.

7.3.2 Imitation Learning Using Neuroevolution of Augmenting Topologies

The two learning techniques presented in the previous section use a common, relatively long learning pipeline. As an alternative approach, we choose to investigate the applicability of a technique which does not require feature selection and PCA for its inputs, but can work directly on the normalized data. The algorithm we choose to apply is the NEAT neuroevolution algorithm [83] augmented with the modifications proposed in the NEAT variant called ANJI [89].

A neuroevolution algorithm creates an artificial neural network by *evolving* the architecture and the weights of a neural network. To improve the evolution process, the NEAT algorithm introduces the concept of the *innovation number*, a historical marking (tagging) of a gene, during the process of evolution. Innovation numbers help in the genetic encoding process and also helps in the protection of speciation. NEAT uses the innovation number of genes during the crossover to identify and track the origin of genes. For every unique structural mutation, a *global innovation number* is attached to the mutated gene. This helps in lining up the genes with similar innovation number during the crossover process.

NEAT evolves the ANNs by the means of crossover and mutations. There are three basic mutations types:

- perturbing the weights of an existing connection in ANN
- adding a new connection between the unconnected weights or connecting a node with itself by the use of a recurrent connection
- splitting the old connection between two neurons by adding a new neuron between them

A fourth type of mutation was introduced for NEAT in one of its implementation called NEAT ANJI [89]. This mutation specifies the rate of deletion for a neural connection between the two nodes. Selected connections are deleted from the stranded genes assuming that the lower weight connections are less influential and are better candidates to be deleted [89].

To apply the NEAT algorithm for modeling the controller, we investigated three different approaches:

a. Complexification - The evolution is initiated with a population of chromosomes having minimal complexity (size of specie), i.e., an ANN with no hidden nodes. The search space moves towards the higher dimensions of an ANNs (by mutation) only when the lower ones provide stagnant outputs.

b. Simplification - The evolution is initiated with a population with complex ANN structures. Over the period of evolution, the size of ANN is pruned till we get a chromosome with optimal output.

c. Blended - The evolution is initiated with a population of mixed properties from both complexification and simplification settings [89].

Variance in the performance of evolved neural network for NEAT depends upon the structure of the initial popoulation. The "simplification" configuration will give good performance if the structure of the initial population is a subset of the required optimal neural network structure. For an unknown solution, it is difficult to predict a good structure for initial population. Hence, a weak prediction of the solution subset, if used for initial population, would lead to an inefficient search space [83].

The issue beforehand for "complexification" configuration is the limited exploration of a topological subset. In "complexification", hill climbing search leads to a local minima (in the search space) and new structures will not be explored. Hence during the evolutionary phase, new structures are sidelined because they have a lower index of fitness. To solve the problem of limited exploration, NEAT uses historical markings to protect speciation: NEAT retains the innovations of new structures. In NEAT, during evolution, a new structure first competes in niche and the new structure is only allowed to compete beyond the niche after it is optimized.

Table 7.5 provides details for the common configuration settings used during the evolutionary phase. Table 7.6 provides the mutation settings used for different configurations for modeling the controller with NEAT.

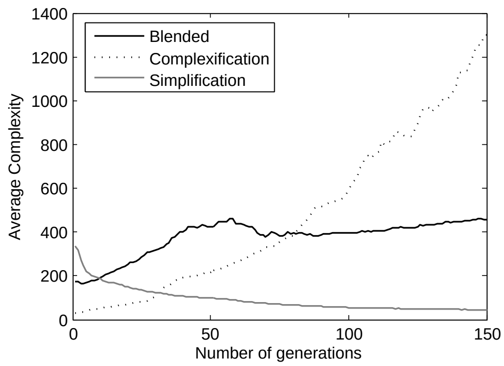
Table 7.5: Common parameter settings for the evolutionary runs

Max. number of generations	150
Population size	500
Remove connection max weight	100
Weight mutation rate	0.8
Weight mutation standard deviation	1.5
Survival rate	0.2
Elitism	True
Speciation threshold	0.2
Roulette selection	Not Used
Topology activation	Sigmoid
Recurrent cycles	Disallowed

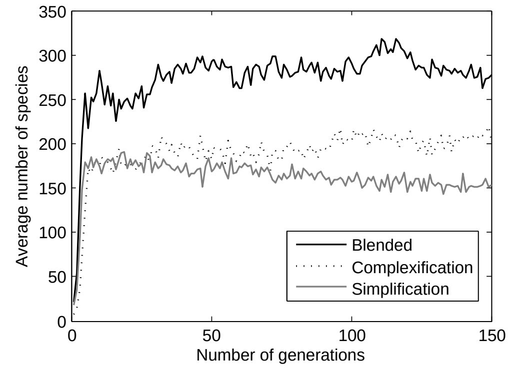
Table 7.6: Parameter settings for different configuration used for evolving controller using NEAT

	Complexification	*Simplification*	*Blended*
Add connection mutation rate	0.03	0.00	0.03
Remove connection mutation rate	0.00	0.04	0.02
Add neuron mutation rate	0.01	0.00	0.01

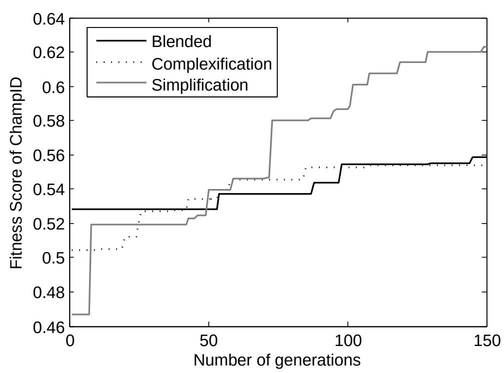
Figure 7.3(a) is the average size of the evolved species during different generations. In comparison with the rest of the evolutionary settings, *simplification* resulted in the smaller sizes of the species over the course of evolution. Similarly, the size of the champ chromosome using *simplification* has the smallest structural size (see Figure 7.3(d)). The confusion matrix of evolved controllers using different configurations is given in Table 7.3.2.



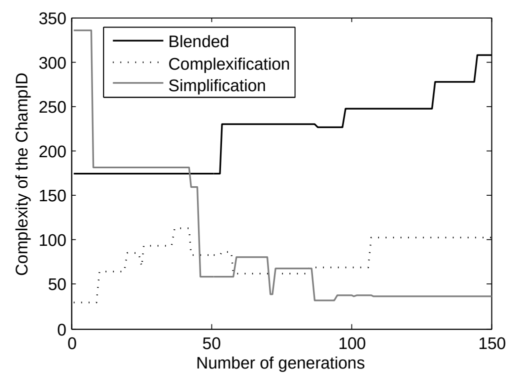
(a) Average complexity (size of species) of population during evolution for different generations



(b) Average number of species over different generations



(c) Fitness of the champion chromosome over the number of generations



(d) Size of the champ chromosome over the number of generations

Figure 7.3: The comparative analysis for different evolutionary techniques used for neuroevolution learning module

Table 7.7: Confusion matrix for controllers evolved different configuration of NEAT

	Complexification		*Simplification*		*Blended*	
	$\bar{C}$	$\bar{D}$	$\bar{C}$	$\bar{D}$	$\bar{C}$	$\bar{D}$
C	99	41	126	69	97	38
D	51	159	24	131	53	162

Figure 7.3(c) give us the comparison of the fitness for the fittest chromosomes that were evolved using complexification, simplification and blended settings. For selecting a neuroevolution based decision-maker, we select the fittest chromosome that was evolved using simplification.

7.4 Evaluation of the Decision-makers

For a situated mobile robot [90], a suitable model for the controller would be able to balance the social costs and mission costs with the dynamically changing environment. For example, in our case, we have considered few features of the environment that are sensed by the mobile robot when it decides to cooperate or defect in a micro-conflict. The density of the crowd should not effect the performance of a controller: a reasonable controller would balance both the social and mission costs while moving in a dense crowd or sparse crowd. For the evaluation of the modeled controllers, we vary the properties of the environment and create four scenarios. The properties that we vary for the scenarios are the density of the crowd and the urgency of mission for the robot. Table 7.8 provides the settings for the four scenarios.

Table 7.8: Parameter settings for scenrios used for evaluation of controllers

	Crowd density	*Mission type*
Scene A	Low	Urgent
Scene B	High	Urgent
Scene C	High	Non-urgent
Scene D	Low	Non-urgent

In Table 7.8, the feature "low density" simulates a market place which is not crowded. Due to this environment setting, the mobile-robot will encounter only one or two opponents in a single micro-conflict. The feature "high density" is the environment setting of an over-crowded marketplace and the mobile-robot will usually encounter more then two opponents in a single micro-conflict. For an "urgent" mission the mobile-robot cannot afford to loose much time in a micro-conflict: it cannot cooperate during all micro-conflicts, while for a "non-urgent" mission, the mobile-robot can cooperate more during micro-conflicts.

For the evaluation, we again collected the data from human operator controlling the mobile-robot in market-place. The decisions were recorded for all of the mobile-robot micro-conflicts in the scenarios mentioned in Table 7.8. It should be noted that the dataset which was used to train various controllers had examples similar to *Scene A* and *Scene B*. A few examples of *Scene D* were incorporated in the training dataset. No example was used from *Scene C* to train the decision-maker controller. Therefore, *Scene C* is used to evaluated the generalisation ability of the modeled decision-makers.

The modeled decision-makers were evaluated using the dataset collected from the human operators. The classification accuracy for different decision-makers is given in Table 7.9. Results show that the classifiers show good classification accuracy for *Scene A* and *Scene B* because most of the examples in the training dataset were included from *Scene A* and *Scene B*. For generalized evaluation, the SVM and Random Forest classifier are tested in an environment created from *Scene C*. Table 7.9 shows that the controllers have low classification accuracy for *Scene C*. During modeling phase, cross-validation was applied for ensemble learning and SVM classifier to avoid overfitting but in this case the low classification accuracy is due to undersampling from *Scene C*, i.e., a bias exists in the training database for *Scene A* and *Scene B*.

For the evolutionary learning, we can see that the classification is comparatively better in all of the four scenarios. One of the reasons for having good accuracy for generalised results is the assumption of not reducing the dimensionality for neuroevolution learning. The other classifiers were trained on datasets after dimensionality reduction. Perhaps some of the features that were pruned during the dimensionality reduction contributed more towards decision making in *Scene C* and hence would had made an effective SVM decision-maker. Due to more generalized results we conclude that neuroevolution learning can be used to imitate controller for decision-making for movement in crowds.

Table 7.9: Classification accuracy for different models

	SceneA	*SceneB*	*SceneC*	*SceneD*
SVM	76%	88.64%	26.09%	48.94%
Random Forests	76%	83.36%	30.43%	48.93%
NEAT	75%	74.42%	59.09%	73.91%

7.5 Validating the imitation learning

Imitation learning of a behavior is successful if the learned behavior can be "mistaken" for the original. As the best judges of the similarity of social behavior are humans, we decided to test the learned behaviors through a user study where we investigated whether human observers can differentiate the behavior of a mixed autonomy robot which resolves micro-conflicts through a human operator from a robot where the micro-conflicts are resolved using the classifier learned as in the previous section.

We have implemented an experimental scenario involving a marketplace in a Middle-Eastern country that has been implemented in the Yaes [62] simulation environment. The area is a narrow space surrounded with shops whose entrances serve as landmarks, as well as internal obstacles. Individual shoppers enter and leave the market and visit the stores. If their trajectories lead them on a course which intersects their personal space or movement cone, they resolve micro-conflicts through a sequence of games and consistent meta-strategy.

In this environment we introduced the presence of a robot of the size of the Big Dog robot [77]. We considered two variations of the robot control:

H: The path planning is controlled through dynamically added waypoints and the D*-Lite algorithm [69], while the micro-conflict behavior is resolved by a human remote operator.

R: The path planning is controlled as in the case of the H robot. The micro-conflict behavior is resolved using the random-forest based controller developed in the previous section.

We created 5 different scenarios with different crowd sizes {20, 30, 40, 50, 100}, and repeated each scenario with an H and an R robot. The resulting 10 scenario runs were recorded as video files 7.1.

Using this set of 10 videos, we conducted an opinion poll using 11 subjects (2 female and 9 male). The subjects were briefed on the fact that the robot is navigating the crowd while balancing the need to reach its destination quickly with the desire to not upset the social norms of politeness. The subjects were informed that in some videos the robot is remotely controlled by a human while in others it acts autonomously. Then, the subjects were requested to identify which video shows a remote controlled (iH) and which an autonomous robot (iR). The 110 responses were then matched with the actual video shown, resulting in the following confusion matrix:

$$\begin{array}{cc} & \begin{array}{cc} iR & iH \end{array} \\ \begin{array}{c} R \\ H \end{array} & \begin{pmatrix} 27 & 29 \\ 28 & 26 \end{pmatrix} \end{array}$$

Essentially, what this result shows is that the answers of the study subjects were not better than random, a fact also confirmed by the debriefing interviews. We conclude that the proposed framework had successfully imitated the human behavior in micro-conflicts and human observers were unable to distinguish between the resolution of micro-conflicts by a human operator and a controller learned by imitating it.

CHAPTER 8
CONCLUSIONS

In order to conform to the social and cultural norms of the environment in which they operate, social robots need to be able to understand and interpret the social behavior of humans. In this dissertation, we presented contributions towards improving human-robot interaction for social robots. The general setup we have considered was one of mobile robots assisting soldiers on a peacekeeping mission. These robots need to understand the local social and cultural conventions, take into account the social norms, but at the same time also accomplish their mission.

The first scenario we considered is a near-future peace-keeping scenario, where a group of soldiers (of different ranks) and a robot interact with the local population in the context of an urban checkpoint near a busy marketplace. We have developed a model using culture sanctioned social metrics (CSSMs) that can be used for learning the impact of actions on social values for a Middle Eastern marketplace. We show that with the use of CSSMs, we can trace the evolution of the social values through individual interactions, rather than a value integrated over populations. We have verified the hypothesis of CSSMs using a survey of the perception of social incurred during the interaction. The survey was administered to people familiar with the Middle Eastern culture. Using the CSSM model, we calibrated the values acquired by the survey and run a series of simulations for modeling the evolution of values over the course of several weeks. The outcome of the social values match well with the intuition and judgment of people with similar cultural background. For the CSSM model, the effect of an action on a CSSM is described by the action-impact function (AIF). We described a method that acquires AIFs from the survey of human respondents who evaluate the impact of a specific sequence of actions on the social values. We used genetic programming to learn AIFs that match the responses of human subjects.

The second scenario we considered was the case of mobile robots moving in a crowd. Such robots need to conform to the same social standards as the human crowd members. Imitating human behavior is a natural choice in these situations - however, not every human behaves in the same way. On the other hand, it is known that humans tend to behave in a consistent way, with their behavior predictable by their social status. We considered a marketplace scenario where humans and the mobile robot perform purposeful movement. With many people moving on intersecting trajectories, the participants occasionally encounter *micro-conflicts*, where they need to balance their desire to move towards their destination (their mission) with the requirements of the social norms of not bumping into strangers or violating their personal space. We model micro-conflicts by a series of two-player games. In the dissertation, we have shown that if a human is using a consistent strategy and is aware of his own social status then it can also infer the social status of its opponent during a micro-conflict. A consistent strategy would minimize the overall social costs as compared to a scenario where the humans use inconsistent strategies (even if those strategies are adaptive). We argued that the correct approach for a robot is not a strategy to avoid all of social costs. Instead, in a micro-conflict, a mobile robot should use a socially consistent strategy that depend of his opponent in the micro-conflict. This would allow the humans to form a mental model of the robot's behavior (a "theory of the robot mind") and adjust their own behavior accordingly.

We developed an alternative approach using the imitation strategy for a mobile robot that would reflect the strategy of humans while moving in crowds. We collected the dataset of the mobile robot controlled by a human operator and used three different supervised learning algorithms (random forest, SVM and neuroevolution) to create a decision maker module. The decision maker module imitates the human operator's behavior in a micro-conflict and hence reflects the same strategy used by humans during micro-conflicts. Results show that the neuroevolution-based decision-maker gives results most closely matching the strategy of humans under scenarios with various crowd density and mission urgency. In addition, we observe that the neuroevolution decision maker generalizes better as it imitates the similar behavior in environments that were not learnt during the user-study.

APPENDIX A
IRB APPROVAL

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Exempt Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Stephen M. Fiore** and Co-PIs if applicable: **Ladislau Boloni**

Date: **September 12, 2012**

Dear Researcher:

On 09/12/2012, the IRB approved the following activity as human participant research that is exempt from regulation:

Type of Review:	Exempt Determination
Project Title:	Culture Specific Models of the Surface Social Values
Investigator:	Stephen M Fiore
IRB Number:	SBE-12-08639
Funding Agency:	DOD/Army/ARL, General Dynamics
Grant Title:	RCTA-H9 Social Cues and Behaviors in HR Collaboration
Research ID:	1054234

This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these changes affect the exempt status of the human research, please contact the IRB. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Signature applied by Patria Davis on 09/12/2012 11:07:26 AM EDT

IRB Coordinator

LIST OF REFERENCES

[1] C. Miller, P. Wu, and H. Funk, "A computational approach to etiquette: Operationalizing Brown and Levinson's politeness model," *IEEE Intelligent Systems*, pp. 28–35, 2008.

[2] C. Miller, P. Wu, V. Vakili, T. Ott, and K. Smith, "Culture, politeness and directive compliance," in *Proc. of the 5th International Conference on Universal Access in Human-Computer Interaction*, pp. 568–577, July 2009.

[3] T. Bosse, C. Jonker, and J. Treur, "Formalisation of Damasio's theory of emotion, feeling and core consciousness," *Consciousness and cognition*, vol. 17, no. 1, pp. 94–113, 2008.

[4] S. Khan, T. Bhatia, S. Parker, and L. Bölöni, "Modeling the interaction between mixed teams of humans and robots and local population for a market patrol task," in *Twenty-Fifth Florida Artificial Intelligence Research Society Conference (FLAIRS-25)*, pp. 50–55, 2012.

[5] S. Khan, T. Bhatia, and L. Bölöni, "Soldiers, robots and local population - modeling cross-cultural values in a peacekeeping scenario," in *21th Behavior Representation in Modeling & Simulation (BRIMS) Conference*, March 2012.

[6] L. Bölöni, "The Spanish Steps flower scam - agent-based modeling of a complex social interaction," in *Proc. of 11th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, pp. 1345–1346, Jun 2012.

[7] M. Haugh and C. Hinze, "A metalinguistic approach to deconstructing the concepts of "face" and "politeness" in Chinese, English and Japanese," *Journal of pragmatics*, vol. 35, no. 10-11, pp. 1581–1611, 2003.

[8] P. Brown and S. Levinson, *Politeness: Some universals in language usage*, vol. 4. Cambridge Univ Pr, 1987.

[9] H. P. Grice, "Logic and conversation," in *Syntax and Semantics, Volume 3: Speech Acts*, pp. 41–58, Academic Press, 1975.

[10] Y. Gu, "Politeness phenomena in modern Chinese," *Journal of Pragmatics*, vol. 14, no. 2, pp. 237–257, 1990.

[11] Y. Matsumoto, “Reexamination of the universality of face: Politeness phenomena in Japanese,” *Journal of Pragmatics*, vol. 12, no. 4, pp. 403–426, 1988.

[12] G. Hofstede, G. Hofstede, and M. Minkov, *Cultures and organizations: software for the mind*. McGraw-Hill Professional, 2010.

[13] A. Vinciarelli, M. Pantic, and H. Bourlard, “Social signal processing: Survey of an emerging domain,” *Image and Vision Computing*, vol. 27, no. 12, pp. 1743–1759, 2009.

[14] A. R. Damasio, *The feeling of what happens: Body and emotion in the making of consciousness*. Houghton Mifflin Harcourt, 1999.

[15] A. Paiva, J. Dias, D. Sobral, and R. Aylett, “Caring for agents and agents that care: Building empathic relations with synthetic agents,” in *Proc. of Autonomous Agents and Multi Agent Systems*, pp. 194–201, 2004.

[16] L. Hall, A. Woods, K. Dautenhahn, D. Sobral, A. Paiva, D. Wolke, and L. Newall, “Designing empathic agents: Adults versus kids,” in *Proc. of Intelligent Tutoring Systems*, pp. 175–196, 2007.

[17] A. Ortony, G. Clore, and A. Collins, *The Cognitive Structure of Emotions*. Cambridge University Press, 1998.

[18] S. Rodrigues, S. Mascarenhas, J. Dias, and A. Paiva, “I can feel it too!: Emergent empathic reactions between synthetic characters,” in *Proc. of Int'l. Conf. on Affective Computing and Intelligent Interaction (ACII)*, 2009.

[19] S. D. Preston and F. De Waal, “Empathy: Its ultimate and proximate bases,” *Behavioral and Brain Sciences*, vol. 25, no. 01, pp. 1–20, 2002.

[20] F. De Vignemont and T. Singer, “The empathic brain: how, when and why?,” *Trends in Cognitive Sciences*, vol. 10, no. 10, pp. 435–441, 2006.

[21] L. Hogg and N. R. Jennings, “Socially intelligent reasoning for autonomous agents,” *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 31, no. 5, pp. 381–393, 2001.

[22] J. C. Harsanyi, “Morality and the theory of rational behavior,” *Social Research*, pp. 623–656, 1977.

[23] L. Bölöni, S. Khan, and S. Arif, “Robots in crowds - being useful while staying out of trouble,” in *Proc. of Intelligent Robotic Systems Workshop (IRS-2013)*, pp. 2–7, 2013.

[24] M. A. Goodrich and A. C. Schultz, “Human-robot interaction: a survey,” *Foundations and Trends in Human-Computer Interaction*, vol. 1, no. 3, pp. 203–275, 2007.

[25] K. Dautenhahn, "Socially intelligent robots: dimensions of human–robot interaction," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 362, no. 1480, pp. 679–704, 2007.

[26] C. Breazeal, "Emotion and sociable humanoid robots," *International Journal of Human-Computer Studies*, vol. 59, no. 1, pp. 119–155, 2003.

[27] M. Moussaid, S. Garnier, G. Theraulaz, and D. Helbing, "Collective information processing and pattern formation in swarms, flocks, and crowds," *Topics in Cognitive Science*, vol. 1, no. 3, pp. 469–497, 2009.

[28] D. Helbing, L. Buzna, A. Johansson, and T. Werner, "Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions," *Transportation Science*, vol. 39, no. 1, pp. 1–24, 2005.

[29] D. Helbing, I. Farkas, and T. Vicsek, "Simulating dynamical features of escape panic," *Nature*, vol. 407, no. 6803, pp. 487–490, 2000.

[30] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Physical review E*, vol. 51, no. 5, p. 4282, 1995.

[31] N. Fridman and G. A. Kaminka, "Towards a cognitive model of crowd behavior based on social comparison theory," in *Proceedings of the National Conference on Artificial Intelligence*, vol. 22, pp. 731–737, 2007.

[32] N. Pelechano, J. M. Allbeck, and N. I. Badler, "Controlling individual agents in high-density crowd simulation," in *Proc. of the ACM/SIGGRAPH/Eurographics Symp. on Computer Animation*, pp. 99–108, 2007.

[33] P. Trautman and A. Krause, "Unfreezing the robot: Navigation in dense, interacting crowds," in *2010 IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pp. 797–803, 2010.

[34] J. A. Kirkland and A. A. Maciejewski, "A simulation of attempts to influence crowd dynamics," in *IEEE Int'l Conf. on Systems, Man and Cybernetics*, vol. 5, pp. 4328–4333, 2003.

[35] A. Chakravarthy and D. Ghose, "Obstacle avoidance in a dynamic environment: A collision cone approach," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 28, no. 5, pp. 562–574, 1998.

[36] S. Guy, J. Chhugani, C. Kim, N. Satish, M. Lin, D. Manocha, and P. Dubey, "Clearpath: highly parallel collision avoidance for multi-agent simulation," in *Proc. of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp. 177–187, ACM, 2009.

[37] J. Van den Berg, M. Lin, and D. Manocha, "Reciprocal velocity obstacles for real-time multi-agent navigation," in *IEEE Int. Conf. on Robotics and Automation*, pp. 1928–1935, 2008.

[38] J. Van Den Berg, S. Guy, M. Lin, and D. Manocha, "Reciprocal n-body collision avoidance," *Robotics Research*, pp. 3–19, 2011.

[39] S. Ge and Y. Cui, "Dynamic motion planning for mobile robots using potential field method," *Autonomous Robots*, vol. 13, no. 3, pp. 207–222, 2002.

[40] L. Huang, "Velocity planning for a mobile robot to track a moving target: a potential field approach," *Robotics and Autonomous Systems*, vol. 57, no. 1, pp. 55–63, 2009.

[41] F. Lamarche and S. Donikian, "Crowd of virtual humans: a new approach for real time navigation in complex and structured environments," in *Computer Graphics Forum*, vol. 23, pp. 509–518, 2004.

[42] S. Waydo and R. Murray, "Vehicle motion planning using stream functions," in *IEEE Int. Conf. on Robotics and Automation*, vol. 2, pp. 2484–2491, IEEE, 2003.

[43] N. Basilico, N. Gatti, and F. Amigoni, "Leader-follower strategies for robotic patrolling in environments with arbitrary topologies," in *Proc. of the 8th Int. Conf. on Autonomous Agents and Multiagent Systems*, pp. 57–64, 2009.

[44] F. Amigoni, N. Gatti, and A. Ippedico, "A game-theoretic approach to determining efficient patrolling strategies for mobile robots," in *Proc. IEEE/WIC/ACM Int. Conf. on Web Intelligence and Intelligent Agent Technology*, vol. 2, pp. 500–503, IEEE, 2008.

[45] P. Paruchuri, J. Pearce, J. Marecki, M. Tambe, F. Ordonez, and S. Kraus, "Playing games for security: an efficient exact algorithm for solving Bayesian Stackelberg games," in *Proc. of the 7th Int. Joint Conf. on Autonomous Agents and Multi-Agent Systems*, pp. 895–902, 2008.

[46] J. Pita, M. Jain, F. Ordóñez, M. Tambe, S. Kraus, and R. Magori-Cohen, "Effective solutions for real-world Stackelberg games: When agents must deal with human uncertainties," in *Proc. of the 8th Int. Conf. on Autonomous Agents and Multiagent Systems*, pp. 369–376, 2009.

[47] J. Pita, M. Tambe, C. Kiekintveld, S. Cullen, and E. Steigerwald, "Guards-game theoretic security allocation on a national scale," in *Proc. of The 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, vol. 3, 2011.

[48] E. Shieh, B. An, R. Yang, M. Tambe, C. Baldwin, J. DiRenzo, B. Maule, and G. Meyer, "PROTECT: A deployed game theoretic system to protect the ports of the united states," in *Proc. of the 11th Int. Conf. on Autonomous Agents and Multiagent Systems*, pp. 13–20, 2012.

[49] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and Autonomous Systems*, vol. 57, no. 5, pp. 469–483, 2009.

[50] A. Billard, S. Calinon, R. Dillmann, S. Schaal, B. Siciliano, and O. Khatib, "Robot programming by demonstration," in *Handbook of Robotics* (B. Siciliano and O. Khatib, eds.), pp. 1371–1394, Springer, 2008.

[51] M. Cakmak, N. DePalma, R. I. Arriaga, and A. L. Thomaz, "Exploiting social partners in robot learning," *Autonomous Robots*, vol. 29, no. 3-4, pp. 309–329, 2010.

[52] C. L. Nehaniv and K. Dautenhahn, *Imitation and social learning in robots, humans and animals: behavioural, social and communicative dimensions*. Cambridge University Press, 2007.

[53] S. Schaal, "Is imitation learning the route to humanoid robots?," *Trends in Cognitive Sciences*, vol. 3, no. 6, pp. 233–242, 1999.

[54] C. Breazeal and B. Scassellati, "Robots that imitate humans," *Trends in Cognitive Sciences*, vol. 6, no. 11, pp. 481–487, 2002.

[55] J. Nakanishi, J. Morimoto, G. Endo, G. Cheng, S. Schaal, and M. Kawato, "Learning from demonstration and adaptation of biped locomotion," *Robotics and Autonomous Systems*, vol. 47, no. 2, pp. 79–91, 2004.

[56] J. Allen and J. Anderson, "An approach to imitation learning for physically heterogeneous robots," in *and Learning Agents Workshop at AAMAS 2010*, Citeseer, 2010.

[57] M. Blokzijl-Zanker and Y. Demiris, "Multi robot learning by demonstration," in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pp. 1207–1208, International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[58] M. Gladwell, *Outliers: The story of success*. Little, Brown and Company, 2008.

[59] P. Brickman and D. Campbell, "Hedonic relativism and planning the good society," in *Adaptation-level theory: a symposium*, pp. 287–302, Academic Press, 1971.

[60] G. Shafer, *A mathematical theory of evidence*. Princeton university press Princeton, NJ, 1976.

[61] R. Yager, "On the Dempster-Shafer framework and new combination rules," *Information sciences*, vol. 41, no. 2, pp. 93–137, 1987.

[62] L. Bölöni and D. Turgut, "Yaes: a modular simulator for mobile networks," in *Proceedings of the 8th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems*, pp. 169–173, ACM, 2005.

[63] J. Koza, *Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems*. Stanford University, Department of Computer Science, 1990.

[64] R. Poli, W. Langdon, and N. McPhee, *A field guide to genetic programming*. Lulu Enterprises Uk Ltd, 2008.

[65] J. Henrich, S. Heine, and A. Norenzayan, "The weirdest people in the world," *Behavioral and Brain Sciences*, vol. 33, no. 2-3, pp. 61–83, 2010.

[66] M. O'Neill and C. Ryan, "Grammatical evolution," *Evolutionary Computation, IEEE Transactions on*, vol. 5, no. 4, pp. 349–358, 2001.

[67] S. Silva and J. Almeida, "GPLAB - a genetic programming toolbox for MATLAB," in *Proceedings of the Nordic MATLAB conference*, pp. 273–278, 2003.

[68] A. Stentz, "The focussed d$\hat{*}$ algorithm for real-time replanning," in *International Joint Conference on Artificial Intelligence*, vol. 14, pp. 1652–1659, 1995.

[69] S. Koenig and M. Likhachev, "Fast replanning for navigation in unknown terrain," *IEEE Transactions on Robotics*, vol. 21, no. 3, pp. 354–363, 2005.

[70] E. Hall and E. Hall, *The hidden dimension*. Anchor Books New York, 1969.

[71] R. Adolphs, "Cognitive neuroscience of human social behaviour," *Nature Reviews Neuroscience*, vol. 4, no. 3, pp. 165–178, 2003.

[72] C. Nass, J. Steuer, and E. R. Tauber, "Computers are social actors," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 72–78, ACM, 1994.

[73] A. G. Sanfey, J. K. Rilling, J. A. Aronson, L. E. Nystrom, and J. D. Cohen, "The neural basis of economic decision-making in the ultimatum game," *Science*, vol. 300, no. 5626, pp. 1755–1758, 2003.

[74] J. Harsanyi, "Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points," *International Journal of Game Theory*, vol. 2, no. 1, pp. 1–23, 1973.

[75] R. Aumann and A. Brandenburger, "Epistemic conditions for nash equilibrium," *Econometrica: Journal of the Econometric Society*, pp. 1161–1180, 1995.

[76] A. Vinciarelli, M. Pantic, and H. Bourlard, "Social signal processing: Survey of an emerging domain," *Image and Vision Computing*, vol. 27, no. 12, pp. 1743–1759, 2009.

[77] M. Raibert, K. Blankespoor, G. Nelson, R. Playter, and the BigDog Team, "Bigdog, the rough–terrain quadruped robot," in *Proc. of Int'l Conf. of Automatic Control*, pp. 10823–10825, 2008.

[78] J. M. Bradshaw, P. J. Feltovich, H. Jung, S. Kulkarni, W. Taysom, and A. Uszok, "Dimensions of adjustable autonomy and mixed-initiative interaction," in *Agents and Computational Autonomy*, pp. 17–39, Springer, 2004.

[79] M. A. Goodrich, D. R. Olsen, J. W. Crandall, and T. J. Palmer, "Experiments in adjustable autonomy," in *Proceedings of IJCAI Workshop on Autonomy, Delegation and Control: Interacting with Intelligent Agents*, pp. 1624–1629, 2001.

[80] S. Koenig and M. Likhachev, "D* lite.," in *Proc. of Nat'l Conf. on Artificial Intelligence*, pp. 476–483, 2002.

[81] S. Khan, S. Arif, and L. Bölöni, "Emulating the consistency of human behavior with an autonomous robot in a market scenario," in *Proc. of Plan, Activity, and Intent Recognition workshop (PAIR-2013)*, pp. 17–23, 2013.

[82] S. Khan, S. Arif, and L. Bölöni, "Towards learning movement in dense crowds for a socially-aware mobile robot," in *Proc. of Adaptive Learning Agents workshop at AAMAS-2014*, 2014.

[83] K. O. Stanley and R. Miikkulainen, "Evolving neural networks through augmenting topologies," *Evolutionary Computation*, vol. 10, no. 2, pp. 99–127, 2002.

[84] I. Jolliffe, *Principal component analysis.* Wiley Online Library, 2005.

[85] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, 2009.

[86] M. E. Mavroforakis and S. Theodoridis, "A geometric approach to support vector machine (SVM) classification," *IEEE Trans. on Neural Networks*, vol. 17, no. 3, pp. 671–682, 2006.

[87] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, p. 27, 2011.

[88] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," tech. rep., Department of Computer Science, National Taiwan University, 2003.

[89] D. James and P. Tucker, "A comparative analysis of simplification and complexification in the evolution of neural network topologies," in *Proc. of Genetic and Evolutionary Computation Conference*, 2004.

[90] J. J. Steil, F. Röthling, R. Haschke, and H. Ritter, "Situated robot learning for multi-modal instruction and imitation of grasping," *Robotics and Autonomous Systems*, vol. 47, no. 2, pp. 129–141, 2004.