A Probabilistic Model of Human-Robot Spatial Interaction Using a Qualitative Trajectory Calculus

Christian Dondrup, Nicola Bellotto, Marc Hanheide
University of Lincoln
Brayford Pool, LN6 7TS Lincoln
United Kingdom
{cdondrup, nbellotto, mhanheide}@lincoln.ac.uk

Abstract

In this paper we propose a probabilistic model for Human-Robot Spatial Interaction (HRSI) using a Qualitative Trajectory Calculus (QTC). In particular, we will build on previous work representing HRSI as a Markov chain of QTC states and evolve this to an approach using a Hidden Markov Model representation. Our model accounts for the invalidity of certain transitions within the QTC to reduce the complexity of the probabilistic model and to ensure state sequences in accordance to this representational framework. We show the appropriateness of our approach by using the probabilistic model to encode different HRSI behaviours observed in a human-robot interaction study and show how the models can be used to classify these behaviours reliably.

Introduction

Human Robot Spatial Interaction (HRSI) is the study of the joint movement of humans and robots through space. It aims to investigate and establish models of the way humans and robots can manage their motion given spatial constraints, how they negotiate, how to move when and where when close to one another, and how to effectively coordinate these movements. Imagine, for instance, the situation of a narrow passage that both a robot and a human aim to pass through. Clearly the robot and the human need to be aware of their mutual goals and have a way to negotiate who goes first. In our work we aim to equip our robot with an understanding of such HRSI situations and enable them to act accordingly.

A large body of research is dedicated to provide answers to these questions on joint spatial behaviour and spatial management in Human-Robot Interaction (HRI). In early works, humans were merely seen as static obstacles (Borenstein and Koren 1989) that had to be avoided by a robot. Then, their dynamic aspects have been taken into account (e.g. Simmons, 1996). In more recent works, robots explicitly plan to move on more "socially acceptable and legible paths" (Sisbot et al. 2007; Yoda and Shiota 1996; Feil-Seifer and Matarić 2011). The term "legible" here refers to the communicative – or interactive – aspects of motions which previously has widely been ignored in robotics research. Looking at spatial behaviour of humans, Ducourant *et al.* (2005) explain that humans also have to consider the actions of others as well, when planning their own actions. Hence, moving around is also about communication and coordination of one's movements – at least in cases where people are walking within a certain vicinity to one another, e.g. entering each other's personal or social spaces (Hall 1968).

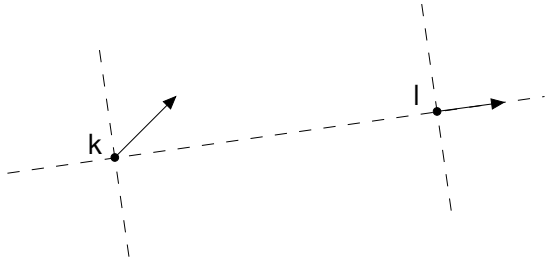

Figure 1: Example of moving points k and l. The respective QTC_C relation is $(- + - 0)$.

The Qualitative Trajectory Calculus

In our previous work (Bellotto, Hanheide, and Van de Weghe 2013; Hanheide, Peters, and Bellotto 2012) we started looking at *qualitative* representations to encode HRSI behaviours and by means of these extract the essence of the joint motion of one robot and one human. We have adopted the Qualitative Trajectory Calculus (QTC) for this purpose which has originally been put forward by Van de Weghe (2004). QTC provides "a language for representing and reasoning about movements of objects in a qualitative framework" (Van de Weghe, Kuijpers, and Bogaert 2005). It allows for a very compact representation of the trajectories of two agents, and – as a calculus – provides a sound mathematical foundation with well-defined rules for composition and qualitative reasoning.

In this paper, we extend our original approach (Hanheide, Peters, and Bellotto 2012) to encode sequences of QTC states in a probabilistic framework as a Hidden Markov Model to both represent and recognise different HRSI behaviours. We present experimental results from a HRSI study to elicit different behaviour in a natural setting and show that our employed variant of QTC states represented

in an HMM is powerful enough to recognise the two behaviours we focused on with very high accuracy.

In analysing human-robot spatial interactions, the use of quantitative (i.e. metrical) representations is often unnecessary or even undesirable, in particular if these representations do not capture the real essence of the observed interaction. In the recent works of Bellotto *et al.* (2012; 2013) and Hanheide *et al.* (2012), therefore, a qualitative representation of interactive human and robot trajectories was proposed based on the Qualitative Trajectory Calculus (QTC) (Van de Weghe 2004). A particular variant of QTC, namely QTC Double-Cross (QTC$_C$), provides enough details to describe, in qualitative terms, complex 2D trajectories of two moving agents with respect to each other. The advantage of such a calculus lies in its simple and intuitive definitions, notwithstanding the rich set of possible operations (e.g. Conceptual Neighbour Diagrams, Composition Tables) to work on temporal sequences of qualitative instances and combinations of multiple interactive agents.

In practice, QTC$_C$ describes the relative motion of two points k and l (Fig. 1) with respect to the reference line connecting them, using a 6-tuple $(q_1\ q_2\ q_3\ q_4\ q_5\ q_6)$, where $q_i \in \{-, 0, +\}$, each one describing a specific qualitative relation:

q_1) movement of k with respect to l
- $-$: k is moving towards l
- 0 : k is stable with respect to l
- $+$: k is moving away from l

q_2) movement of l with respect to k

as above, swapping k and l

q_3) relative speed of k with respect to l
- $-$: k is slower than l
- 0 : k has the same speed of l
- $+$: k is faster than l

q_4) movement of k with respect to $\overrightarrow{k\,l}$
- $-$: k is moving to the left side of $\overrightarrow{k\,l}$
- 0 : k is moving along $\overrightarrow{k\,l}$
- $+$: k is moving to the right side of $\overrightarrow{k\,l}$

q_5) movement of l with respect to $\overrightarrow{l\,k}$

as above, swapping k and l

q_6) minimum absolute angle of k, α_k, with respect to $\overrightarrow{k\,l}$ at time t
- $-$: $\alpha_k < \alpha_l$
- 0 : $\alpha_k = \alpha_l$
- $+$: $\alpha_k > \alpha_l$

A simplified version of QTC$_C$ is often adopted excluding the terms q_3 and q_6 about speed and angle respectively. Therefore, the relative motion of two points at time t can be expressed by a 4-elements state descriptor $(q_1\ q_2\ q_4\ q_5)$, yielding a total number of $3^4 = 81$ possible states. The state $(-+-0)$, for example, means "k moves towards l; l moves away from k; k moves on the left-hand side of $\overrightarrow{k\,l}$; l moves along $\overrightarrow{l\,k}$" (see Fig. 1).

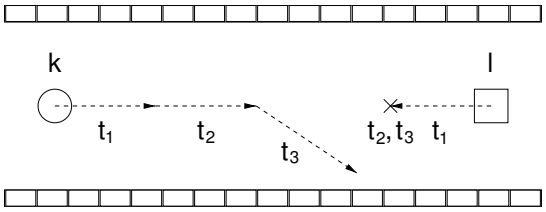

Figure 2: Example of spatial interaction where the robot (agent l) stops to let the person (agent k) pass on its left.

Combined in temporal sequences, QTC states can be used to represent different scenarios of spatial interaction between two agents. Consider the following example, illustrated also in Fig. 2: a person and a robot move towards each other along a narrow corridor; when close enough, the robot stops to let the person pass on its left-hand side. If k is the person and l the robot, this situation can be described in QTC$_C$ by the following temporal sequence where t_i represents the temporal interval during which the QTC$_C$ state holds:

$$(-\ -\ 0\ 0)_{t_1} \rightsquigarrow (-\ 0\ 0\ 0)_{t_2} \rightsquigarrow (-\ 0\ +\ 0)_{t_3} \qquad (1)$$

The possible transitions between QTC states can also be represented using a Conceptual Neighbourhood Diagram (CND), which is shown in Fig. 3 for the simplified QTC$_C$. Transitions between neighbouring states are characterized by changes in one or more of the q_i symbols: the larger the number of symbols simultaneously changed, the higher the "conceptual distance" between two states. Note however that the CND is not a fully connected graph, i.e. only a small set of inter-state transitions are allowed. In practice, a CND restricts the number of possible transitions and helps to reduce the complexity in constructing temporal sequences of QTC states and relative Hidden Markov Models, as discussed in the following section. See (Van de Weghe 2004) for more details on CNDs.

A Probabilistic Model of QTC sequences

In order to encode and facilitate automatic recognition and reasoning about joint Human-Robot Spatial behaviour, the sequence of QTC states needs to be represented. In our previous work (Hanheide, Peters, and Bellotto 2012) we first proposed a Markov model as a suitable framework for a compact representation of (prototypical) behaviour, enabling both recognition of currently observed behaviour and prediction of future behaviour to control the robot appropriately. This model has now been taken a step further to also constrain the state transitions according to the CND to ensure valid QTC sequences. In addition to the respective QTC states (e.g. 81 states in the simplified version of QTC$_C$) dedicated states for the beginning and the end of a sequence are added. In our model, transitions from the start state and into the end state are allowed for all QTC states in addition to the transitions stemming from the QTC-specific CND.

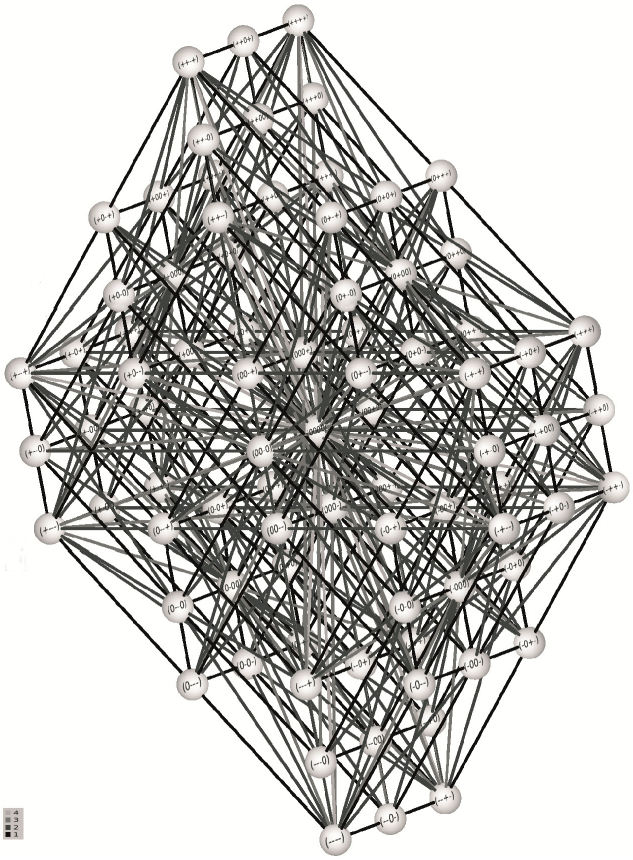

Figure 3: CND of the simplified QTC$_C$ (from Delafontaine, 2011). Note that, due to the original formulation of CNDs, there are no direct transitions between some of the states, e.g. $(-\ 0\ +\ -)$ and $(-\ +\ 0-)$.

Also, this model is now evolved into a Hidden Markov Model (HMM) (Fink 2008) that models the automatically recognised QTC states as emissions, allowing for uncertainty in the actual recognition process. This allows us to deal with QTC state classification errors that arise from the discretisation of actual robot and human movement into the respective qualitative states. We have initially modelled the "correct" emissions (e.g. the QTC state $(-\ 0\ +\ -)$ actually emits $(-\ 0\ +\ -)$) to occur with 95% probability and allow the model to account for classification errors with 5%.

This general approach applicable to the different variants of QTC has been implemented to represent HRSI encoded in the simplified version of QTC$_C$. Hence, our HMM for QTC$_C$ sequences features a total of 83 states (the 81 QTC states plus the start and the end state), with respective 83 emissions. There are a total of 1250 possible transitions, 1088 from the CND of QTC$_C$ and $2 \cdot 81 = 162$ from and to the start and end states, respectively.

To represent different HRSI behaviours, the probabilistic model needs to be trained from actual observed data. For each different behaviour to be represented, a separate HMM is trained using Baum-Welch training (Fink 2008) (Expectation Maximisation) to obtain the appropriate transition and emission probabilities for the respective behaviour. In the initial model before training, transitions that are *valid* according to the CND are modelled as equally probable (uniform distribution). In order to overcome the problem of a lack of sufficient amounts of training data and unobserved transitions therein, we allow for pseudo transitions with a very low probability of $P_{pt} = 1e^{-10}$. Fig. 8 depicts a pruned HMM for one behaviour that will be discussed in detail in the following.

Representing Real-World Interactions

A first approach of using simplified QTC$_C$ on sensor data was conducted using a motion capture system, tracking the movements of a robot and a human in a confined space. In the following we will give an overview of the experimental design, the data processing, and the final representation in form of the previously described HMM.

Experiment

The cover story for this experiment was a hypothetical restaurant scenario with the participant as a waiter and a robotic co-worker. As part of their job the participants had to pick up drinks from the kitchen counter and serve them at one of two tables while the robot was taking orders from the supposed guests and delivered them to the kitchen counter. This allowed us to send the participants from the kitchen counter (start position) to one of two tables (goals) and back while the robot was also continuously moving between the start position and either of the two goals. This restaurant scenario was well suited to elicit incidental human-robot encounters and allowed us to observe spontaneous human movements when encountering the robot while they were trying to fulfil their task as efficient as possible.

The purpose of this study was to find typical spatial movements of human participants when confronted with certain kinds of robot behaviour. The two behaviours shown by the robot in this pilot study were: *a) non-stopping*: only regarding the human as a static obstacle and trying to reach the goal as fast and efficient (in forms of path planning) as possible, and *b) stopping*: gradually reducing its speed and coming to a complete halt before entering the personal space (Hall 1968) of the participant. The latter behaviour was created by combining the findings of Lichtenthäler *et al.* (2013), i.e. stopping to let people pass, and Basili *et al.* (2012), i.e. slowing down to let people pass, which both are normally used in human-robot or human-human path crossing scenarios. Even though we did not expect many path crossing situations we decided to use these two behaviours, i.e. stopping and non-stopping, because both produce very similar trajectories which only differ in speed and distance when approaching a person. Using these behaviours we are trying to find differences in human reactions regarding the two mentioned factors, i.e. speed and distance, by analysing subtle movements using the motion capture system. If this first study shows encouraging results, we will also deviate from this simple approach and evaluate different path planning strategies.

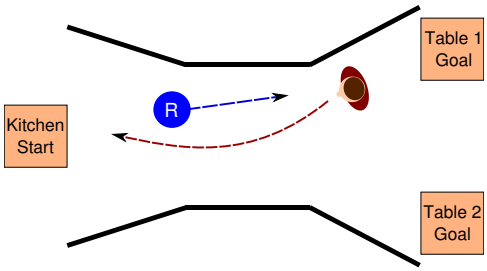

Figure 4: Head on encounter. The robot tries to reach the table while the human is trying to reach the kitchen counter. This image and Fig. 5 also show the experimental set-up with the kitchen counter (start) on the left and the two tables (goals) on the right. The artificial corridor is represented by black lines. The circle labelled "R" represents the robot and the human-like figure represents the human.

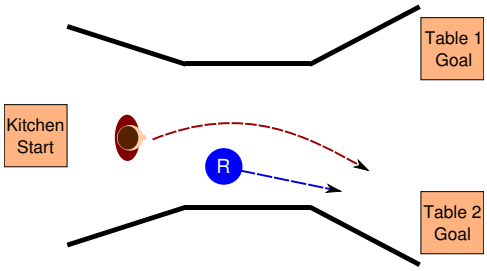

Figure 5: Overtaking scenario. According to observations during the experiment, the human was much faster than the robot resulting in overtaking situations where both agents tried to reach the same goal.

As mentioned, both of these behaviours are very similar regarding the actual path the robot takes so we were able to pool them and disregard the robot's specific behaviour throughout the rest of this paper. Indeed, we will show in the following sections that we are allowed to disregard these sub-divisions because both behaviours produce very similar QTC_C state sequences in similar situations. Also, the division into these two behaviours is not necessary to show the functionalities of our HMM based modelling approach and is only described for the sake of completeness.

Experiment Set-Up The actual experiment was conducted in a large gymnasium like room fitted with 12 motion capture and two video cameras (see Fig. 6). In this pilot study we chose to use a motion capture system, in addition to the above mentioned reason of finding subtle human reactions, to create a precise baseline for future experiments and to overcome one of the limitations of the robot, namely the lack of sensors in the back. Hence, to model the movement of the participant behind the robot we needed these external sensors.

The experimental set-up consisted of a start and two target positions between which the participant and the robot moved. These positions were marked by two large boxes, resembling tables, and a barstool to act as the kitchen counter.

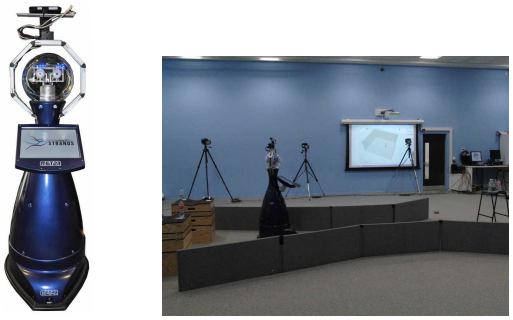

Figure 6: The SCITOS G5 robot and the physical experimental set-up. The robot has a hight of $1.72m$ and a diameter of $\sim 61cm$. This picture shows the back of the robot. The experimental set-up had a size of $5m^2$. The "tables" are on the left and the "kitchen counter" can be seen on the right of the image.

In addition, a narrow passage — in form of an artificial corridor built out of 50cm high polystyrene boards — between the start and the two target positions was created (see Fig. 4 and 5). This corridor was used to elicit close encounters between the human and the robot while still allowing the motion capture system to reliably track the position of both agents.

The robot used was a SCITOS G5 with a human-machine-interaction superstructure produced by MetraLabs (see Fig. 6). It was working completely autonomously during the experiment and showed the two aforementioned behaviours of stopping and non-stopping which were switched at random when arriving at the start position (kitchen counter). Both of these behaviours, as mentioned above, are very similar in forms of path planning which means that they both resulted in mostly straight movements directly towards the goal.

The robot was fitted with 7 motion capture markers to get a precise position estimate and to also be able to represent the physical dimensions of the robot itself. Since the robot as a whole is a rigid structure, this set-up allowed us to track 2 Degrees Of Freedom (DOF), i.e. movement in x and y direction. The human participant was fitted with 3 markers on a hat and 4/5 markers on two shoulder pads which allowed us to track 6 DOF, i.e. pitch, yaw, and roll for the head and upper body, including the overall movement in x and y direction, of the human participant. Note that, for the purpose of this paper, only the letter was necessary to create a full QTC representation of human motion.

During our study we recorded 14 participants, of which 9 have a computer science background. They are all employees or students at the university, and only two had previous experience with robots. Each participant interacted with the robot in the described way for 6 minutes.[1]

[1]The raw data set containing the recorded motion capture sequences is publicly available on our git repository: https: // github . com / LCAS / data

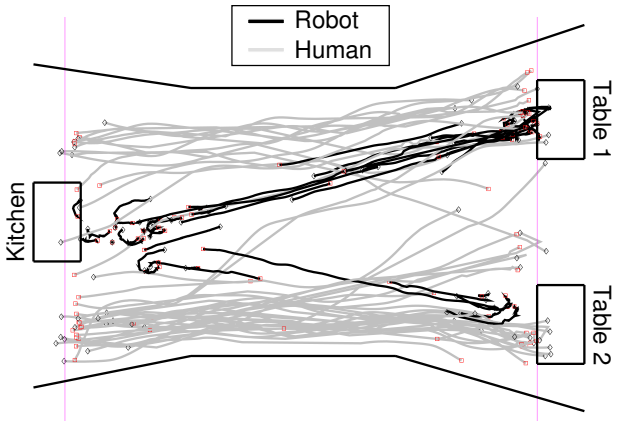

Figure 7: The recorded trajectories of one of the participants (grey = human, black = robot).The rough position of the corridor walls and the furniture is also depicted.

HMM Based QTC$_C$ Representation

To create a QTC$_C$ representation of the gathered data we had to select trajectories for further analysis from the raw data set. In this study we were mainly interested in the close encounters of the robot and the human in the corridor to evaluate how a robot's behaviour influences the movements of a human in a confined space. To generate these trajectories from our data we defined a virtual cut-off line at either side of the corridor and only took the points in between those two lines into account. If a participant changed her/his direction of movement before reaching either of those two lines, the trajectory was cut at the maximum or the minimum x value respectively to ensure that a single trajectory only contains human movements in one direction along the x-axis. This resulted in approximately 50 trajectories per person and 729 trajectories in total.

In the final trajectories, according to the definition of QTC, both the robot and the human are only represented by one point in space. This also allowed the creation of more robust trajectories because, due to occlusion, not all markers were always visible. Representing the human and robot only as one single point in space each enabled us to use the visible markers to still infer their position and therefore compensate for possible occlusion. Fig. 7 shows the paths of one of the participants and the robot as recorded by the motion capture system and processed in the above mentioned way. The selection process for the robot's trajectories does not adhere to the same constraints as the trajectory selection for the human participant but are just chosen according to the timestamps of the participant's trajectory.

Out of the created set of trajectories, we selected two types of HRSI situations for further analysis. These are the *head-on* (see Fig. 4) and *overtaking* (see Fig. 5) encounters between the robot and the participant. We decided to select these two classes because the head-on and overtaking scenarios are the two most common encounters between robot and human in our data set where the trajectories of each agent might influence the trajectory of the other and vice versa. All trajectories in which the robot and the human did not encounter each other in the corridor, because the robot was stationary while turning at one of the target positions, or the trajectories in which the robot and human tried to reach different goals and did not come in to close vicinity of each other were disregarded because of the lack of possible influence of the agents on each other.

Our two chosen classes of encounters would normally be divided further into the above mentioned sub-classes according to the behaviour the robot showed at that specific time. For the QTC$_C$ representation and analysis these two subclasses are pooled together because the actual trajectories only differ in their duration and just marginally in terms of spatial movement. Therefore, we chose to only use the two classes head-on and overtaking as a whole and disregarded the division according to the behaviour. This is possible, because as a qualitative representation, QTC$_C$ is not able to represent time in general but just temporal sequences. In the following sections we will show that the two behaviours are not significantly different in their QTC$_C$ representation.

Manual annotation of the whole dataset yielded 122 head-on and 85 overtaking trajectories for all participants. These were automatically converted into QTC$_C$ state sequences using two different means of noise reduction.[2] The actual trajectory was smoothed by averaging over the x and y positions for $0.1s$ and $0.3s$ respectively. The z value cannot be represented in QTC$_C$ and was therefore disregarded. Due to the fact that actual sensor data will most likely never produce 0 states in QTC$_C$, the distance to $\overrightarrow{k\,l}$ and the two $\perp\ \overrightarrow{k\,l}$ lines was thresholded with either $1cm$, $5cm$, or $10cm$ distance from either of those lines. The resulting state chain is then post-processed to remove equal, adjacent states.

HMM Modelling of QTC$_C$ Sequences

The initial HMM used to model the QTC$_C$ state sequences is created as explained previously and afterwards the QTC$_C$ state sequences created from the motion capture data set are used to train the HMM using Baum-Welch training. The generated HMM can then be used to visualise the observed states, classify HRSI encounters, and predict the outcome of those encounters.

A typical QTC$_C$ graph created by a HMM trained on the recoded data is a directed graph and is composed of states which represent the actual QTC$_C$ states and directed edges symbolising the transitions and their probabilities to go from one state to another. The particular graph shown in Fig. 8 shows all the state sequences created for the head-on encounters using a $0.3s$ averaging interval over the (x, y) positions and a $5cm$ accuracy threshold regarding the distance to either $\overrightarrow{k\,l}$ or the two $\perp\ \overrightarrow{k\,l}$ lines. For the sake of readability and space all transition probabilities below $p = 0.15$ have been pruned and the states without incoming or outgoing edges have been removed. As a result of the pruning, most of the transition probabilities do not sum up to $p = 1.0$ but the graph still reflects the most probable transitions and states during such an encounter.

[2]The motion capture system operated with an error of approximately $1.5mm$ to $2.5mm$.

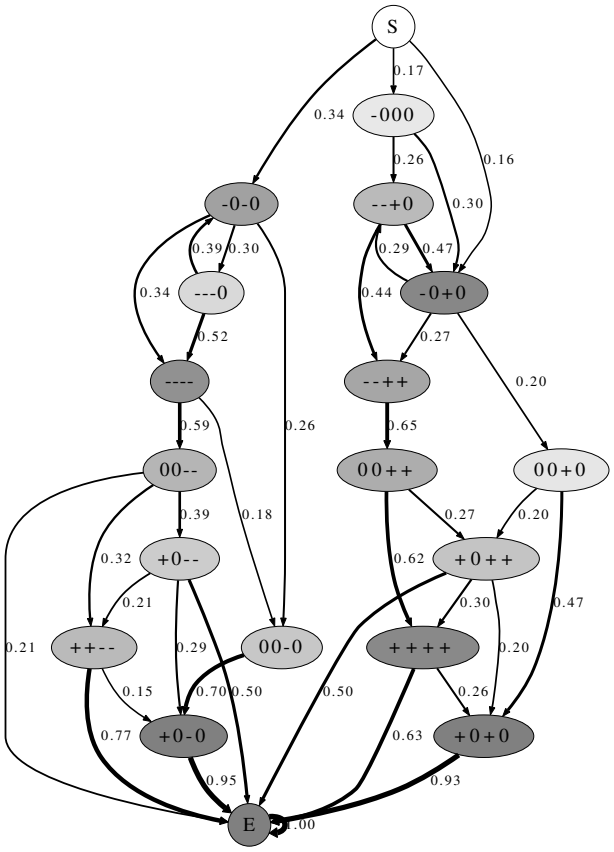

Figure 8: QTC$_C$ states for all head-on encounters created by the HMM representation. Edge width represents the transition probability. The colour of the nodes represents the a-priori probability of that specific state (from white = 0.0, e.g. "S", to dark grey = 1.0, e.g. "E"). Representation has been created with an accuracy of $5cm$ and an averaging over $0.3s$. All transition probabilities below 0.15 have been pruned from the graph, only highlighting the most probable paths within our model. Due to the pruning the transition probabilities in the graph do not sum up to 1.0.

One of the most probable paths (disregarding the start and end state (S, E) which are artificially inserted during post-processing to represent uncertainty about states that precede the start and succeed the end of our evaluated trajectory) through the graph shown in Fig. 8 could be interpreted as follows [3]:

$(-\ 0\ -\ 0)$ The human moves towards the robot while the robot is stationary or turning on the spot. The human already starts to circumvent the robot by moving to the left of $\overrightarrow{k\,l}$.

$(-\ -\ -\ 0)$ The human and robot approach each other and the human still moves to the left of $\overrightarrow{k\,l}$. This might not mean that he/she is circumventing the robot but just that the angle of the connecting line has changed. The human might be moving straight on.

$(-\ -\ -\ -)$ Human and robot are approaching each other and are moving to the left of $\overrightarrow{k\,l}$ and $\overrightarrow{l\,k}$ respectively. As mentioned, this might not mean that they are circumventing but move straight on with a changed angle of the connecting line which would fit the observed general behaviour of participants and robot.

$(0\ 0\ -\ -)$ The movement along $\perp\ \overrightarrow{k\,l}$ is the actual passing-by action where both agents are moving in opposite directions while being "shoulder to shoulder".

$(+\ 0\ -\ -)$ Both agents have passed each other and the human – if the previous assumption that the human is actively circumventing the robot was correct – is correcting her/his path by moving behind the robot.

The described state sequence is very similar to the prototypical head-on encounter depicted in Fig. 4.

HMM Based QTC$_C$ Classification

After demonstrating that the HMM approach can be used to automatically create and model valid representations of QTC$_C$ state chains, we also used the resulting HMM as a classifier to show that the data is correctly modelled and to be able to identify similar situations in our data set. In order to show the accuracy of this classification we used the ground truth information, obtained by manually sorting and categorising the recorded trajectories, as a basis and compared them against each other. The HMM was trained in the same way as described in the previous sections, using the HMM initially created from the CND and training it on the actual data using a Baum-Welch training.

We then defined two hypotheses which we are trying to prove using this HMM based QTC$_C$ classification:

Hypothesis 1: The two classes, i.e. head-on and overtaking, can be distinguished by the trained HMM and are significantly different from each other.

Hypothesis 2: The two sub-classes, i.e. stopping and non-stopping, cannot be distinguished by the trained HMM and are therefore not significantly different from each other. They mainly differ in the time domain which cannot be modelled by QTC$_C$.

To determine if our assumptions are correct we used k-fold cross validation with $k = 5$ over 10 iterations, resulting in 50 iterations in total. During each of the single validation processes we used the k-fold training sub-set to train a new HMM to model this particular sub-set of the data and used the remaining samples as a validation or test set. The test set was used to test the generated HMM by producing a log-likelihood value representing the probability of producing the given test set with this particular HMM based on its transition and emission probabilities. As a control, the same was done with a HMM trained on the other class given the

[3] k represents the human and l the robot. As a result $(q_1\ q_2\ q_4\ q_5)$ are defined as: q_1 = human moves towards or away from robot, q_2 = robot moves towards or away from human, q_4 = human moves to the left or right of $\overrightarrow{k\,l}$, q_5 = robot moves to the left or right of $\overrightarrow{l\,k}$.

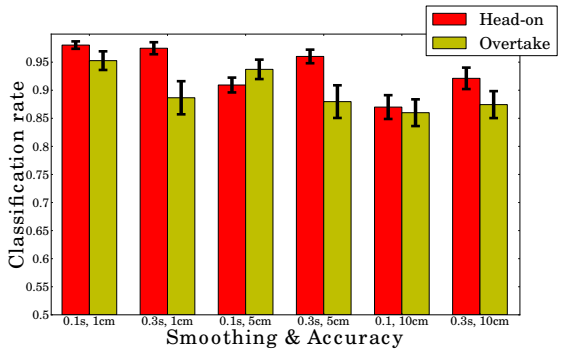

Figure 9: The classification rates for the different smoothing times and accuracy thresholds. Red (dark grey) represents the head-on class, yellow (light grey) represents the overtaking class. The error bars represent the 95% confidence interval. Note: To better visualise the small differences between the results, the bottom line does not represent 0 but $p = 0.5$ (the null hypothesis for our two-class problem).

same test set, also generating a log-likelihood value representing the production probability of the test set for the control HMM. A single sample in the test set was considered correctly classified if the HMM modelling the correct class produced a higher log-likelihood than the control HMM modelling the other class. The number of correctly classified state sequences in the test set were normalised thereby creating a classification rate for every single k-fold iteration.

Afterwards, a normal distribution was fitted over these 50 classification rates to estimate the mean, variance, and the according 95% confidence intervals. The results of this validation process will be described in the following section.

Results

Using the above mentioned validation process, comparing our two classes of HRSI encounters, we achieved classification rates from $\mu = 0.8700$ to $\mu = 0.9804$ for the head-on class and $\mu = 0.8600$ to $\mu = 0.9527$ for the overtaking class (see Fig. 9). The best classification rate was produced by the lowest filter settings with a (x, y) position averaging over $0.1s$ and an accuracy threshold of $1cm$. We were able to achieve a classification rate of $\mu = .9804 \pm .0066$ with a standard deviation of $\sigma = .0193$ for the head-on case and a classification rate of $\mu = .9527 \pm .0166$ with a standard deviation of $\sigma = .0488$ for the overtaking case. The specified mean intervals represent the 95% confidence intervals on said mean value and the results can therefore be interpreted as being significantly different from the null hypothesis. This means that the two classes are distinguishable because the confidence interval is well above $\mu = 0.5$.

The general classification accuracy seems to be loosely correlated to the spatial accuracy used to determine which states count as 0-states in the QTC$_C$ representation, therefore indicating that this is the largest factor of data loss in our automatic construction of QTC$_C$ states. Nevertheless, even the $10cm$ accuracy results are still significantly different from the null hypothesis and therefore support our first hypothesis that the HMM approach is well suited to model QTC$_C$ state sequences and classify new data.

An in-class comparison of the two previously mentioned robot behaviours, i.e. stopping and non-stopping, did not yield significant classification results. If we, for example, have a look at the best case in the two class classification of head-on and overtaking[4] obtained with a $0.1s$ smoothing and an accuracy of $1cm$ we achieve classification results, for the in-class comparison, of $\mu = 0.5946 \pm 0.0452$ for the head-on non-stopping, $\mu = 0.6560 \pm 0.0419$ for the head-on stopping, $\mu = 0.4077 \pm 0.0511$ for the overtake non-stopping, and $\mu = 0.5107 \pm 0.0515$ for the overtake stopping scenario. This indicates that the two behaviours are too similar regarding the actual relative movement of the two agents because they mainly differ in the time domain, as we suggested in our second hypothesis. These findings justify the pooling of the data and to only consider the head-on and overtaking classes as a whole for our evaluation process.

Conclusion and Discussion

In this paper we built on previous work to improve QTC$_C$ state sequence modelling by evolving the previously proposed Markov Chain representation into a Hidden Markov Model based approach. The initial HMM is built upon CND based transition and emission probabilities to create valid representations of the observed temporal sequences. Afterwards, we employed a Baum-Welch training to fit the HMM to the actual observed data from a real user study. We were able to correctly model the recorded motion capture data, to automatically create valid QTC$_C$ representations, and to use these models as classifiers for newly observed sequences. By proving our two hypotheses we showed that for our two class HRSI examples of encounters in a narrow corridor we are able to reliably distinguish and classify them.

Besides the shown advantages of representing relative motion of a robotic and human agent as QTC$_C$ states, we also found a few limitations to this approach of data representation.

Normally, one of the limitations of QTC$_C$ is that actual sensor data does not coincide with the constraints represented in a CND. In the mentioned motion capture study we encountered up to 521 illegal transitions which indicates that raw sensor data is not suitable to create QTC$_C$ state sequences without post-processing. This however, was solved by using our proposed HMM based modelling adhering to the constraints defined in the CND, only producing valid state transitions.

A major limitation is that three important factors of HRSI, namely speed, acceleration, and distance, cannot be represented by simplified QTC$_C$. While full QTC$_C$ is able to represent relative speeds it is neither possible to represent the velocity nor acceleration of the robot or the human. It is also impossible to determine the distance between the two agents and therefore to adhere to the social distances defined by Hall (1968) when creating robot behaviour based on QTC$_C$ states or evaluating if these distances were kept when analysing data. Therefore, QTC$_C$ is not very well suited to

[4]The results for the other cases were not significantly different.

make statements about *comfort*, *naturalness*, and *sociability*, as defined by Kruse *et al.* (2013), of a given HRSI encounter. One possible starting point for incorporating HRSI principles into QTC_C could be the representation of Hall's personal distance as another factor in the QTC tuple, e.g. $0 =$ personal distance, $- =$ closer, $+ =$ further away. The same might also be possible for velocity and acceleration in close vicinity of humans.

Another limitation of QTC_C is the impossibility to infer which agent executes the actual circumvention action in the head-on scenario. As described previously when interpreting the graph in Fig. 8, we are not sure if the human, the robot, or both are circumventing each other. We just know that the human started the action but we do not know if the robot participated or not. This could eventually be countered by using the full QTC_C approach including the relative angles. Even then, it might not be possible to make reliable statements about that and it would also complicate the graph and deprive it of some of its generalisation abilities.

Future Work

As an immediate follow-up to this evaluation of the HMM based modelling approach we would not only like to use it as a classifier but also to predict the behaviour of the human. In the preliminary work of Bellotto *et al.* (2012; 2013), predefined QTC sequences were already encoded into a symbolic reasoner to generate motion commands for a robot that approached and/or avoided a nearby user. We hope to enhance this form of QTC reasoning by being able to produce the most probable path in our HMM representation, infer the unfolding of a specific HRSI encounter, and act appropriately.

In future work, we will try to tackle some of the limitations described in the previous section, e.g. by introducing a qualitative representation of the distance between the two agents and other forms of qualitative representation which have yet to be explored. This will be necessary to enable QTC_C based behaviour creation to adhere to common HRSI principles and constraints.

Both the behaviour generation and the enriching of QTC will be evaluated in subsequent user studies, exploring different forms of motion tracking. This will also indicate if the internal sensors of the robot are sufficient to infer motion patterns and reason on their outcome.

Acknowledgments

The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement No. 600623, STRANDS.

References

Basili, P.; Sağlam, M.; Kruse, T.; Huber, M.; Kirsch, A.; and Glasauer, S. 2012. Strategies of locomotor collision avoidance. *Gait & Posture*.

Bellotto, N.; Hanheide, M.; and Van de Weghe, N. 2013. Qualitative design and implementation of human-robot spatial interactions. In *Proc. of Int. Conf. on Social Robotics (ICSR)*, 331–340.

Bellotto, N. 2012. Robot control based on qualitative representation of human trajectories. In *AAAI Spring Symposium – Designing Intelligent Robots: Reintegrating AI*. TR SS-12-02.

Borenstein, J., and Koren, Y. 1989. Real-time obstacle avoidance for fact mobile robots. *IEEE Transactions on Systems, Man and Cybernetics* 19(5):1179–1187.

Delafontaine, M. 2011. *Modelling and Analysing Moving Objects and Travelling Subjects: Bridging theory and practice*. Ph.D. Dissertation, Department of Geography, Ghent University.

Ducourant, T.; Vieilledent, S.; Kerlirzin, Y.; and Berthoz, A. 2005. Timing and distance characteristics of interpersonal coordination during locomotion. *Neuroscience Letters* 389(1):6–11.

Feil-Seifer, D. J., and Matarić, M. J. 2011. People-Aware Navigation For Goal-Oriented Behavior Involving a Human Partner. In *Proceedings of the International Conference on Development and Learning*.

Fink, G. A. 2008. *Markov Models for Pattern Recognition*. Springer-Verlag Berlin Heidelberg.

Hall, E. T. 1968. Proxemics. *Current Anthropology* 9(2/3):83–108.

Hanheide, M.; Peters, A.; and Bellotto, N. 2012. Analysis of human-robot spatial behaviour applying a qualitative trajectory calculus. In *RO-MAN, 2012 IEEE*, 689–694. IEEE.

Kruse, T.; Pandey, A. K.; Alami, R.; and Kirsch, A. 2013. Human-aware robot navigation: A survey. *Robotics and Autonomous Systems* 61(12):1726–1743.

Lichtenthäler, C.; Peters, A.; Griffiths, S.; and Kirsch, A. 2013. Social navigation-identifying robot navigation patterns in a path crossing scenario. *ICSR. Springer*.

Simmons, R. 1996. The curvature-velocity method for local obstacle avoidance. In *IEEE International Conference on Robotics and Automation*, volume 4, 3375–3382. Minneapolis, Minnesota: IEEE.

Sisbot, E.; Marin-Urias, L.; Alami, R.; and Simeon, T. 2007. A Human Aware Mobile Robot Motion Planner. *IEEE Transactions on Robotics* 23(5):874–883.

Van de Weghe, N.; Kuijpers, B.; and Bogaert, P. 2005. A Qualitative Trajectory Calculus and the Composition of its Relations. *GeoSpatial* 281:60–76.

Van de Weghe, N. 2004. *Representing and Reasoning about Moving Objects: A Qualitative Approach*. Ph.D. Dissertation, Ghent University.

Yoda, M., and Shiota, Y. 1996. Analysis of human avoidance motion for application to robot. In *Proceedings 5th IEEE International Workshop on Robot and Human Communication. RO-MAN'96 TSUKUBA*, 65–70. IEEE.