Tractable Probabilistic Models for Intention Recognition Based on Expert Knowledge

Oliver C. Schrempf, David Albrecht, and Uwe D. Hanebeck

Abstract— **Intention recognition is an important topic in human-robot cooperation that can be tackled using probabilistic model-based methods. A popular instance of such methods are Bayesian networks where the dependencies between random variables are modeled by means of a directed graph. Bayesian networks are very efficient for treating networks with conditionally independent parts. Unfortunately, such independence sometimes has to be constructed by introducing so called hidden variables with an intractably large state space. An example are human actions which depend on human intentions and on other human actions. Our goal in this paper is to find models for intention-action mapping with a reduced state space in order to allow for tractable on-line evaluation. We present a systematic derivation of the reduced model and experimental results of recognizing the intention of a real human in a virtual environment.**

I. INTRODUCTION

The development of humanoid robots is a field of increasing interest in robotics today [1]. An example for this is the humanoid robot system ARMAR III of the Collaborative Research Center 588 [2] "Humanoid Robots - Learning and Cooperating Multimodal Robots".

Creating humanoid robots is not limited to the resemblance to the human body. An important topic is also the question how a humanoid robot can cooperate with humans. This problem is addressed in robotics as human-robot cooperation, human-robot interaction, or social interactive robots [3]. Considering social aspects as a key issue in human-robot cooperation has been pointed out by Breazeal in [4]. A major challenge inherited from human-human cooperation lies in predicting the cooperating counterpart. By anthropomorphising a robot, this becomes relevant in human-robot cooperation as well [5]. Humans expect the robot to "read their intentions" because they are used to this behavior from human counterparts. A robotic system, on the other hand, can gain a lot of information considering the human's intentions since it allows for a new quality of high-level task planning. An example would be proactive behavior on the robot side based on systematically minimizing the uncertainty about the user's intention [6].

The problem of intention recognition is tackled in various fields of human-machine interaction [7]. In robotics, intention recognition has been addressed in several contexts like mobile robotics [8] or programming by demonstration [9].

Since humans are used to deal with uncertainty and ambiguities in everyday life it comes quite natural to use probabilistic methods for inferring the user's intention. A very popular example of probabilistic methods in intention recognition is the Lumière project [10], where the goals of a user were inferred from his actions in office computer applications. For this purpose Bayesian networks where applied. Bayesian networks are a class of graph based probabilistic models exploiting causal dependencies in order to reduce the complexity of large systems [11].

In [12], we have developed a theory of a generic multi-level model for intention recognition based on Bayesian networks. This model provides a level of human intentions a level for human actions and a level for resulting sensor measurements.

In the construction of a Bayesian network for intention recognition, it is often assumed, that actions depending on intentions are conditionally independent in order to have simple models. Unfortunately, this assumption is not always valid and results in incorrect inferences. An exact model that accounts for the dependencies on the other hand may be very complex in its structure and is not tractable for available inference algorithms. A common way to deal with such complex models is to introduce so called hidden nodes, which model the dependencies via joint state spaces [13]. A drawback of this approach is that the joint state space to be considered may be extremely large which also leads to intractability. Hence, methods are needed for reducing the number of states in the hidden variable while maintaining acceptable results in the inference. In general it's not possible to define the cardinality and the adequate conditional probabilities of hidden nodes without investing a lot of time and effort. Hence, several publications address the problem. If not previously known, the hidden nodes of a network have to be discovered [14] [15]. The cardinality of the hidden nodes has to be determined [16] [17] and the parameters of the conditional densities have to be learned from data [18] [19]. A general approach to state space reduction is given in [20], but the focus is more on discrete states resulting from quantization of continuous variables.

In this paper we present an alternative way to construct a Bayesian network for intention recognition without using algorithms for learning the structure or the conditional probabilities. This is accomplished by introducing a hidden variable with a tractable state space, based on knowledge of an expert.

For the evaluation of the proposed models we have developed a experimental platform based on a virtual environment using extended range telepresence techniques [21]. This allows for easily repeatable experiments focussing on real humans without struggling to much with typical hardware problems.

The remainder of the paper is organized as follows. We first give a detailed problem formulation in Sec. II followed

(a) Joint Probability Model (b) Decoupled Model

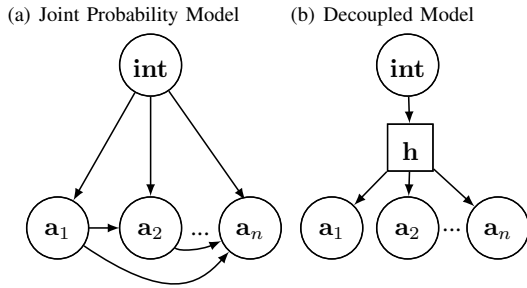Fig. 1. Bayesian network models for intention/action relations. Human intentions are modeled by the discrete random variable **int** and actions are modeled by binary random variables $\mathbf{a}_i$. (a) shows the model for the joint probability. (b) shows the decoupled model using a hidden random variable **h**.

by the derivation of the reduced model in Sec. III. After that we will present our experimental platform in Sec. IV and some results in Sec. V. Conclusions are given in Sec. VI.

II. PROBLEM FORMULATION

When addressing the problem of intention recognition, we can usually depend on the knowledge of human experts providing models of the relationship between intentions and human actions. Such models can easily be given as a mapping from a certain intention to a set of actions that are triggered by this intention. Based on such a mapping we want to build a probabilistic model, where the intention is modeled as a discrete random variable **int** having as many states as there are intentions to be represented. The actions are modeled by means of binary random variables $\mathbf{a}_i$, $i = 1, \ldots, n$, where n is the number of observable actions. Uncertainty in the mapping can be represented by means of a conditional probability $P(a_1, \ldots, a_n | int)$.

Unfortunately, the actions cannot by be modeled independently given the intention, which would allow for independent models $P(a_i | int)$. Since there are dependencies between the actions, we have to consider the joint state space of the involved random variables. A joint state space model of the considered problem is given in Fig. 1(a) by means of a Bayesian network. A Bayesian network is a graphical model, where random variables are represented by nodes in a directed acyclic graph (DAG). Each edge in this graph represents a dependency between the variables, which can be modeled by means of conditional probabilities. The joint state space model of the network in Fig. 1(a) is hence

$$P(int)P(a_1|int)P(a_2|a_1, int) \ldots P(a_n|int, a_1, \ldots, a_{n-1}) .$$

This representation has the drawback that it contains loops in the DAG which renders it unusable for standard message passing techniques that are applied for inferring the intention from observed actions. Although, there are successful applications of loopy belief propagation for Turbo Codes, it is not clear if this algorithm converges for arbitrary networks containing loops [22].

A common approach for getting rid of the loops is to introduce a so called hidden variable as depicted in Fig. 1(b). This leads to a decoupling of the network and a less complex model

$$P(int)P(h|int)P(a_1|h) \ldots P(a_n|h) .$$

The price to be paid for this less complex model is that in the naive application of this approach, the full joint state space has to be modeled as states of the hidden variable. In our case, this leads to a hidden variable with $|\mathbf{int}| \cdot 2^n$ states, where $|\mathbf{int}|$ is the number of known intentions and n is the number of observable actions. In the following, we will refer to this solution as the *exact* solution.

Dealing with such extremely large state spaces is an intractable problem for real world applications. Hence, we have to find an approximate solution which gets by with fewer states in the hidden variable while yielding results comparable to the exact solution, i.e., the inference of the intention from observed actions using the approximate solution must give comparable results to the inference using the exact solution.

For the remainder of this paper we will write the conditional models as $P(\cdot | \cdot)$ in order to emphasize that we deal with discrete random variables for which conditional probabilities can be given by means of probability tables.

III. DECOUPLED MODEL

We will now derive an approximate solution to the decoupling problem given in the previous section that needs remarkably fewer states in the hidden variable while maintining a performance comparable to the exact solution. The derivation of the approximate solution will be accompanied by an example. The example is very simple, in order to keep it comprehensible. Furthermore, we want to start with the exact solution which can only be given for small problems and derive the approximate solution from it.

Example III.1 We consider a case with three known intentions int_1, int_2, int_3 and three possible actions a_1, a_2, a_3 a human can perform. From a human expert, we get a mapping that says which actions can be triggered by which intention. In this example we assume the following mapping:

$$\begin{array}{rcl} int_1 & \to & a_1, a_2, \\ int_2 & \to & a_2, \\ int_3 & \to & a_3. \end{array} \qquad (1)$$

Note, that the first intention can trigger two actions. There is no assumption, that these actions have to occur simultaneously. Note further, that the first and the second intention can trigger the same action a_2.

A. Exact Decoupling

Decoupling the model by adding a hidden variable **h** as introduced in Sec. II leads to a representation of the joint state space $(\mathbf{int}, \mathbf{a}_1, \ldots, \mathbf{a}_n)$ as states of the hidden variable. This makes $|\mathbf{int}| * 2^n$ states in **h**, where each state represents a possible combination of occurrence of $\mathbf{a_i}$ together with a particular intention.

The combination of intentions and joint action states can also be modeled in the conditional probability table $P(a_1, \ldots, a_n | int)$. Hence, the $|int|$ part can be omitted in **h**, which leads to a variable with 2^n states. This modification is no approximation since the mapping from intention to joint actions is just moved from the random variable **h** into the conditional probability $P(h|int)$ and we still have the *exact* solution. Nevertheless, 2^n is still a very large number of states making the model intractable for non-toy examples.

Example III.2 (cont'd) We introduce a random variable **h** with 2^n states, where $n = 3$. Each state h_i, $i = 0, \ldots, 7$ of **h** is

mapped to a corresponding joint action state. Each joint event $a_1 \wedge a_2 \wedge a_3$ is expressed by the binary number $a_1 a_2 a_3$, e.g the occurrence of all actions is indicated by 111. This leads to the following mapping:

$$
\begin{array}{llll}
h_0 \leftarrow 000 & h_1 \leftarrow 001 & h_2 \leftarrow 010 & h_3 \leftarrow 011 \\
h_4 \leftarrow 100 & h_5 \leftarrow 101 & h_6 \leftarrow 110 & h_7 \leftarrow 111
\end{array}
$$

Having introduced the hidden variable $\mathbf{h}$, we can now derive the parameters of the conditional probability table $P(h|int)$ from the intention to actions mapping provided by the human expert.

Before proceeding, we have to consider a case that is not covered by the expert's mapping – the human may have a certain intention in his head, but is not performing any action at all. This may be the case, when the human is in a transition phase where he/she moves from one place to another where the relevant action can be performed. In order to account for this "idle" state, we have to consider it in the probability table $P(h|int)$.

Example III.3 (cont'd) The conditional table for the joint state space for the probabilities $P(h|int)$ can be derived from the mapping of intentions to actions (1) provided by the expert.

In this example, we assume a ratio between the idle state h_0 and the other states of 1:9, which results in the conditional probability $P(h_0|int) = 0.1$. Hence, the remaining conditional probabilities must sum up to 0.9.

$P(h|int)$	$P(h_0)$	$P(h_1)$	$P(h_2)$	$P(h_3)$
int_1	1/10	0	3/10	0
int_2	1/10	0	9/10	0
int_3	1/10	9/10	0	0

$P(h|int)$	$P(h_4)$	$P(h_5)$	$P(h_6)$	$P(h_7)$
int_1	3/10	0	3/10	0
int_2	0	0	0	0
int_3	0	0	0	0

Note, that for the first intention, the probability over actions is distributed over the exclusive and the simultaneous occurrence of actions a_1 and a_2.

B. Reduction

Modeling the full joint state space in $\mathbf{h}$ is intractable for large numbers of observable actions. Hence, we are looking for a reduced state space. The key assumption that can be made in order to gain this reduction is, that in general actions occur exclusively. This means, that people usually perform only one of the actions provided in the mapping at a time. This assumption can be translated into a logical OR-combination of the actions given a certain intention.

Hence, we can reduce the state space of $\mathbf{h}$ by combining all states where a certain event a_i is present, which is equivalent to the OR-combination of all relevant events. This leads to a random variable $\mathbf{h}'$ with $n + 1$ states. The probability $P(h'|int)$ can be calculated from $P(h|int)$ using probability calculus. Here, we make use of the rule, that for two events a and b

$$P(a \vee b) = P(a) + P(b) - P(ab) \ . \tag{2}$$

Example III.4 (cont'd) We introduce a random variable $\mathbf{h}'$ with four states. The states represent the OR-combination of the relevant states from $\mathbf{h}$

$$
\begin{array}{lll}
h'_0 & \leftarrow & h_0 \\
h'_1 & \leftarrow & h_4 \vee h_5 \vee h_6 \vee h_7 \\
h'_2 & \leftarrow & h_2 \vee h_3 \vee h_6 \vee h_7 \\
h'_3 & \leftarrow & h_1 \vee h_3 \vee h_5 \vee h_7.
\end{array}
$$

The state representing that no action has occurred stays untouched as $\mathbf{h}' = h'_0$.

The probability $P(h'_1|int)$ can be calculated from $P(h|int)$ by using probability calculus as

$$
\begin{array}{rcl}
P(h'_1|int) & = & P(h_4 \vee h_5 \vee h_6 \vee h_7 \mid int) \\
& = & P(h_4 \mid int) + P(h_5 \mid int) \\
& + & P(h_6 \mid int) + P(h_7 \mid int)
\end{array}
$$

We made here use of the fact, that the states of $\mathbf{h}$ always occur exclusively. Hence, the last term in (2) is zero. Similar we have for h'_2 and h'_3

$$
\begin{array}{rcl}
P(h'_2|int) & = & P(h_2 \mid int) + P(h_3 \mid int) \\
& + & P(h_6 \mid int) + P(h_7 \mid int) \\
P(h'_3|int) & = & P(h_1 \mid int) + P(h_3 \mid int) \\
& + & P(h_5 \mid int) + P(h_7 \mid int)
\end{array}
$$

The corresponding conditional probabilities $P(h'|int)$ can now be fully determined from $P(h|int)$. The reduced conditional probabilities have to be normalized to sum up to 0.9, which results in

$P(h'|int)$	$P(h'_0)$	$P(h'_1)$	$P(h'_2)$	$P(h'_3)$
int_1	1/10	9/20	9/20	0
int_2	1/10	0	9/10	0
int_3	1/10	0	0	9/10

We now have to consider the models describing how the actions a_i depend on the hidden variable. In the case of the exact model this is straight-forward. The probability tables $P(a_i|h)$ contain only 1 and 0 entries since they provide a direct mapping from the joint state space to the variables involved. We will show this in the next example.

In order to calculate the conditional probabilities $P(a_i|h')$ from $P(a_i|h)$ we have to remember, that the states of $\mathbf{h}'$ are a logical OR-combination of states from $\mathbf{h}$. Hence, we can add up all distributions over $\mathbf{a}_i$ where the condition is relevant. This is shown in the next example.

Example III.5 (cont'd) The probability tables $P(a_i|h)$ for the exact model are given by

$P(a_i|h)$	$P(\bar{a}_1)$	$P(a_1)$	$P(\bar{a}_2)$	$P(a_2)$	$P(\bar{a}_3)$	$P(a_3)$
h_0	1	0	1	0	1	0
h_1	1	0	1	0	0	1
h_2	1	0	0	1	1	0
h_3	1	0	0	1	0	1
h_4	0	1	1	0	1	0
h_5	0	1	1	0	0	1
h_6	0	1	0	1	1	0
h_7	0	1	0	1	0	1

Since h'_0 is identical to h_0 nothing has to be changed here. For the other states of $\mathbf{h}'$ we have to consider the OR-combination. Since

$$h'_1 = h_4 \vee h_5 \vee h_6 \vee h_7$$

we can calculate the conditional probabilities as

$$P(a_i|h'_1) = P(a_i|h_4) + P(a_i|h_5) + P(a_i|h_6) + P(a_i|h_7)$$

and normalize. This leads to a uniform distribution for a_2 and a_3.

The same holds for the other states of $\mathbf{h}'$ where the conditional probabilities can be calculated according to

$$P(a_i|h'_2) = P(a_i|h_2) + P(a_i|h_3) + P(a_i|h_6) + P(a_i|h_7)$$

and

$$P(a_i|h'_2) = P(a_i|h_1) + P(a_i|h_3) + P(a_i|h_5) + P(a_i|h_7) \ .$$

The resulting conditional probability tables $P(a_i|h')$ are hence,

$P(a_i\|h')$	$P(\bar{a}_1)$	$P(a_1)$	$P(\bar{a}_2)$	$P(a_2)$	$P(\bar{a}_3)$	$P(a_3)$
h'_0	1	0	1	0	1	0
h'_1	0	1	1/2	1/2	1/2	1/2
h'_2	1/2	1/2	0	1	1/2	1/2
h'_3	1/2	1/2	1/2	1/2	0	1

C. Comparing the Models

We will now compare the approximate model to the exact model derived from the mapping of actions to intentions provided by a human expert. For this purpose, we compare the estimated distribution over intentions given a set of observed actions.

Example III.6 (cont'd) We start in this example with the observation of the actions as provided from the expert's mapping (1). The estimates using the exact model are given in Tab. I and the estimates using the approximate model are given in Tab. II.

TABLE I

INTENTION ESTIMATE USING THE EXACT MODEL

a_1	a_2	a_3	$P(int_1)$	$P(int_2)$	$P(int_3)$
0	0	0	$0.\bar{3}$	$0.\bar{3}$	$0.\bar{3}$
1	0	0	1	0	0
0	1	0	0.25	0.75	0
0	0	1	0	0	1
1	1	0	1	0	0

TABLE II

INTENTION ESTIMATE USING THE APPROXIMATE MODEL

a_1	a_2	a_3	$P(int_1)$	$P(int_2)$	$P(int_3)$
0	0	0	$0.\bar{3}$	$0.\bar{3}$	$0.\bar{3}$
1	0	0	1	0	0
0	1	0	$0.\bar{3}$	$0.\bar{6}$	0
0	0	1	0	0	1
1	1	0	0.5	0.5	0

It can be seen, that for the occurrence of no action as well as for the occurrence of actions that are mapped to only one intention, the estimate is identical in both models. In the case of action a_2 which is mapped to two intentions, the uncertainty in the estimate using the approximate model is slightly increased. In the case of the simultaneous occurrence of actions a_1 and a_2 this is more apparent since the estimate reflects our assumption in the approximate case, that actions in general occur exclusively.

Tab. III presents the estimate using the approximate model given observed actions combination that were not provided in the expert's mapping. In the exact model, these observations yield an undefined estimate, i.e., all probabilities in the estimate are zero.

TABLE III

INTENTION ESTIMATE USING THE APPROXIMATE MODEL

a_1	a_2	a_3	$P(int_1)$	$P(int_2)$	$P(int_3)$
0	1	1	0.2	0.4	0.4
1	0	1	$0.\bar{3}$	0	$0.\bar{6}$
1	1	1	$0.\bar{3}$	$0.\bar{3}$	$0.\bar{3}$

The example shows, that the approximate model yields results similar to the results of the exact model. Furthermore, it allows for estimates based on observations that are undefined in the exact model.

Deriving the approximate model from the exact model as presented in this section is only possible for small examples, since the exact model in general cannot be given for a large number of actions. Hence, it is desired to derive the approximate model directly from the expert's mapping. Taking a closer look at the approximate model shows, that this is very easy, since the states of $\mathbf{h}'$ (except for h'_0) represent the occurrence of one action. This allows for easily distributing the probability mass over the actions given in the mapping. At first the probability $P(h_0)$ has to be determined. We then use the fact, that each state h'_i with $i > 0$ is associated with exactly one action a_i. Given a certain intention int_i we have to find $P(h'|int_i)$. A value 1 is assigned to every $P(h'_i|int_i)$ which is mapped to this intention by the expert. Then the values are normalized to sum up to $1 - P(h_0)$. Hence, building the approximate model is straight-forward. This is the method, which we will apply in Sec. V.

D. Further Improvement

In practice, we cannot say that an action occured with 100% certainty, since this occurrence is also estimated from sensor observations available to the robot. Due to sensor noise and ambiguities in the action recognition models, the action estimates in general bear uncertainties. These uncertainties have a large impact on the intention estimate, which is mainly caused by the uniform distributions in the conditional probabilities $P(a_i|h')$.

There is room for further improvement of the estimate by adding some knowledge about the simultaneous occurrence of the actions to the models by means of probability distribution over the relevant states of $\mathbf{h}$. The idea is shown in the next example.

Example III.7 (cont'd) The knowledge that it is unlikely (but not impossible) that relevant actions occur together can be modeled by a probability distribution like

$P(\mathbf{h} = h_4)$	$P(\mathbf{h} = h_5)$	$P(\mathbf{h} = h_6)$	$P(\mathbf{h} = h_7)$
5/10	2/10	2/10	1/10

for h'_1, and similar for h'_2 and h'_3. This leads to a weighted sum in the derivation of the reduced models $P(a_i|h')$ as can be seen in the next table.

$P(a_i\|h')$	$P(\bar{a}_1)$	$P(a_1)$	$P(\bar{a}_2)$	$P(a_2)$	$P(\bar{a}_3)$	$P(a_3)$
h'_0	1	0	1	0	1	0
h'_1	0	1	7/10	3/10	7/10	3/10
h'_2	7/10	3/10	0	1	7/10	3/10
h'_3	7/10	3/10	7/10	3/10	0	1

IV. EXPERIMENTAL PLATFORM

In order to evaluate the proposed reduced models for intention recognition we have built a setup using extended range telepresence techniques together with a virtual user environment. The benefit of this setup is, that we can focus on the evaluation of the models while minimizing the hardware effort. Nevertheless, the models are built in a way, that they can be applied to a real robot at any time, given the necessary action recognizers are available.

To support natural interaction with the virtual environment the user's head and hand motion are tracked and reproduced in the virtual environment as described in [21]. The illustration of the virtual world is transferred back and presented to

the user via a head mounted display. Furthermore a motion compression technique [23] allows the user to move around in a larger target environment while the size of the user environment is limited. By using a data-glove the user is able to interact with objects. In order to gain a more natural appearance of the environment a physics simulation engine is applied.

The environment we use for testing our algorithms resembles a kitchen scenario, where the human operator can carry around items, open or close a dishwasher, a fridge or cupboards. In addition he can cooperate with a simulated robot. Furthermore we can easily repeat test runs. Fig. 2 shows the human operator using the telepresence system and Fig. 3 shows the virtual environment from the perspective of the human user who is carrying an apple. Note, that only the hand of the user is modelled but not the full arm, so the hand and the apple are hovering in front of the user.

V. EXPERIMENTAL RESULTS

We will now present some results for intention recognition using a model with a reduced hidden variable as presented above. The relevant intentions to be recognized are

Load dishwasher	Wash dishes
Cook	Lay table
Get drink from robot	Get object from robot

and they are mutual exclusive. With regard to the latter we consider the following 16 actions:

Open/Close dishwasher	Open/Close cupboard
Open/Close fridge	Take food
Take item from cupboard	Put item into cupboard
Put item into sink	Place item onto sink
Place item onto table	Put food onto table
Place item onto workspace	Take item from workspace
Put item into dishwasher	Take item out of dishwasher
Take item from sink	Take item out of sink

The mapping from these intentions to these actions is provided by an expert. From this mapping, we derived the reduced intention recognition model. The exact model for this configuration would have $2^{16} = 65536$ states in the hidden variable. The approximate model needs only 17 states in the hidden variable.

Fig. 2. User working in the virtual kitchen using the extended range telepresence system.

Fig. 3. View of the human operator carrying an apple using the telepresence system.

Since the actions are hidden and cannot be observed directly, we have to extend the model in order to estimate the occurrence of actions from directly observable data. This extension contains continuous as well as discrete random variables an considers temporal dependencies. Hence, we have a hybrid dynamic Bayesian network (HDBN) [24] [25].

Observable inputs to the HDBN are the position of the human hand, the information about an existing hand grasp, as well as the positions of the objects in the virtual kitchen. This includes plates, mugs, and food objects like an apple, as well as positions of door handles. Furthermore, the positions of relevant places in the kitchen like workspaces, a table, the sink and several cupboards are considered. We further observe the distance between the human and the robot in order to distinguish intentions dealing with the robot directly, like "Get drink from robot" from other intentions like "Cook". The full network for recognizing six intentions via sixteen known actions from directly observable data contains more than 300 nodes, which is quite large.

In an experiment, a human performed common actions in the virtual kitchen. First he started cooking. While the food was at the cooker he laid the table. After that he fetched the food from the cooker and moved it to the table. Then he put the dishes into the dishwasher. Finally he wanted to get a drink from the robot. Fig. 4 shows the estimate of the intentions over time. The upper six plots show the estimated probability of the according intention. The plot in the lowest line shows the entropy of the estimate over time as a measure of uncertainty.

VI. CONCLUSIONS

In this paper, we presented probabilistic models for recognizing human intentions based on the observation of actions. We derived a tractable model by reducing the state space of hidden variables while maintaining the relevant parts of the model.

The derivation of the reduced model mainly relies on the assumption that actions in general do not occur simultaneously. Our observations have shown, that the assumption is valid in many cases. From this assumption, the reduction is derived in a mathematically rigorous way.
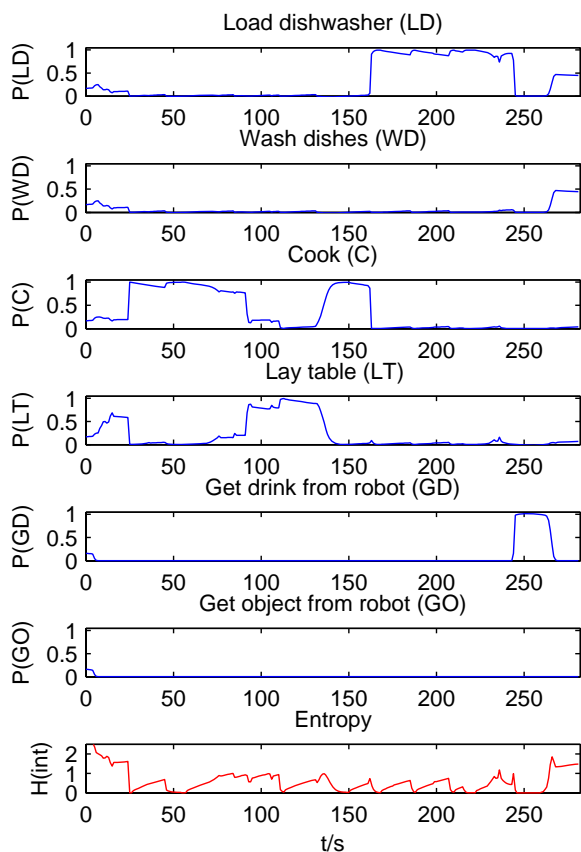
Fig. 4. Estimated probabilities for each possible intention over time. In addition the entropy is presented as a measurement for the uncertainty in the estimate. The intentions of the observed human were cook (C), lay the table (LT), cook, load dishwasher (LD), and get drink from robot (GD)

A key benefit of the proposed reduced model is, that it can be derived directly from an expert's mapping of intentions to actions. It is not necessary to take a detour via the full joint state model.

We have incorporated the proposed approach in a larger model, where actions have to be estimated themselves from available sensor data. Such a model can be applied directly to a real robot.

We have implemented an evaluation platform based on telepresence techniques for testing the proposed model in a virtual environment with real humans. The model shows a good performance and is hence, ready to be tested on a humanoid.

REFERENCES

[1] B. Adams, C. Breazeal, R. Brooks, and B. Scassellati, "Humanoid Robots: A New Kind of Tool," *Intelligent Systems and Their Applications, IEEE [see also IEEE Intelligent Systems]*, vol. 15, no. 4, pp. 25–31, 2000.

[2] R. Becher, P. Steinhaus, and R. Dillmann, "The Collaborative Research Center 588: Humanoid Robots – Learning and Cooperating Multimodal Robots," in *Proceedings of Humanoids 2003, Karlsruhe, Germany*, 2003.

[3] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A Survey of Socially Interactive Robots," *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 143–166, Mar. 2003. [Online]. Available: http://www.sciencedirect.com/science/article/B6V16-47XG102-1/2/aae36740a7e349fe44fdae6bd229dabf

[4] C. Breazeal and B. Scassellati, "How to Build Robots that make Friends and Influence People," in *Proceedings of the IROS*, 1999.

[5] K. Dautenhahn, "Robots as Social Actors: AURORA and The Case of Autism," in *Proceedings of the Third International Cognitive Technology Conference, CT99, San Francisco*, August 1999.

[6] A. Schmid, O. Schrempf, U. Hanebeck, and H. Wörn, "A Novel Approach to Proactive Human-Robot Cooperation," in *Proceedings of the 14th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2005), Nashville, USA*, August 2005.

[7] N. Lesh, C. Rich, and C. Sidner, "Using Plan Recognition in Human-Computer Collaboration," in *Proceedings of the Seventh International Conference on User Modeling, Banff, Canada*, 1999, pp. 23–32.

[8] K. A. Tahboub, "Intelligent Human-Machine Interaction Based on Dynamic Bayesian Networks Probabilistic Intention Recognition," *Journal of Intelligent Robotics Systems*, vol. 45, no. 1, pp. 31–52, 2006.

[9] R. Zöllner, O. Rogalla, R. Dillmann, and M. Zöllner, "Understanding User's Intention: Programming Fine Manipulation Tasks by Demonstration," in *Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Lausanne, Switzerland*, 2002.

[10] E. Horvitz, J. Breese, D. Heckerman, D. Hovel, and K. Rommelse, "The Lumière Project: Bayesian User Modelling for Inferring the Goals and Needs of Software Users," in *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, AUAI. Morgan Kaufman, jul 1998, pp. 256–265.

[11] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.

[12] O. C. Schrempf and U. D. Hanebeck, "A Generic Model for Estimating User-Intentions in Human-Robot Cooperation," in *Proceedings of the 2nd International Conference on Informatics in Control, Automation and Robotics, ICINCO 05*, 2005.

[13] J. Pearl, "Fusion, propagation, and structuring in belief networks," *Artificial Intelligence*, vol. 29, pp. 241–288, 1986.

[14] X. B. N. Friedman and D. Koller, "Discovering the Hidden Structure of Complex Dynamic Systems," in *Proceedings of the 15th Conference on Uncertainty in Articicial Intelligence —UAI1999*. San Francisco: Morgan Kaufmann, 1999, pp. 91–100. [Online]. Available: http://www.cs.stanford.edu/˜xb/uai99/

[15] G. Elidan, N. Lotner, N. Friedman, and D. Koller, "Discovering Hidden Variables: A Structure-Based Approach," in *NIPS*, 2000, pp. 479–485. [Online]. Available: citeseer.ist.psu.edu/elidan01discovering.html

[16] J. Martin and K. VanLehn, "Discrete Factor Analysis: Learning Hidden Variables in Bayesian networks." Department of Computer Science, University of Pittsburgh, Tech. Rep., 1995.

[17] G. Elidan and N. Friedman, "Learning the dimensionality of hidden variables," in *Uncertainty in Artificial Intelligence: Proceedings of the Seventheenth Conference*, 2001, pp. 144–151.

[18] C.-K. Kwoh and D. Gilles, "Using Hidden Nodes in Bayesian Networks," *Artificial Intelligence*, vol. 88, pp. 1–38, 1996.

[19] S. Ramachandran and R. J. Mooney, "Theory Refinement for Bayesian Networks with Hidden Variables," in *Proceedings of the Fifteenth Internation Conference on Machine Learning*, 1998.

[20] M. P. Wellman and C.-L. Liu, "State-Space Abstraction for Anytime Evaluation of Probabilistic Networks." in *Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence, Seattle, Washington, USA*, 1994, pp. 567–574.

[21] H. Groenda, F. Nowak, P. Rößler, and U. D. Hanebeck, "Telepresence Techniques for Controlling Avatar Motion in First Person Games." in *Intelligent Technologies for Interactive Entertainment, First International Conference, INTETAIN 2005, Madonna di Campiglio, Italy,*, 2005, pp. 44–53.

[22] K. P. Murphy, Y. Weiss, and M. I. Jordan, "Loopy Belief Propagation for Approximate Inference: An Empirical Study." in *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence, Stockholm, Sweden*, 1999, pp. 467–475.

[23] N. Nitzsche, U. D. Hanebeck, and G. Schmidt, "Motion Compression for Telepresent Walking in Large Target Environments," *Presence: Teleoperators & Virtual Environments*, vol. 13, no. 1, pp. 44–60, 2004. [Online]. Available: http://www.mitpressjournals.org/doi/abs/10.1162/105474604774048225

[24] O. C. Schrempf and U. D. Hanebeck, "A New Approach for Hybrid Bayesian Networks Using Full Densities," in *Proceedings of 6th Workshop on Computer Science and Information Technologies, CSIT 2004*, Budapest, Hungary, 2004. [Online]. Available: http://isas.uka.de/downloads/csit2004.pdf

[25] ——, "Evaluation of Hybrid Bayesian Networks using Analytical Density Representations," in *Proceedings of the 16th IFAC World Congress, IFAC 2005*, Prague, Czech Republic, 2005.