2006 Special issue

Goals and means in action observation: A computational approach

Raymond H. Cuijpers [a,*], Hein T. van Schie [a], Mathieu Koppen [a], Wolfram Erlhagen [b], Harold Bekkering [a]

[a] *Nijmegen Institute for Cognition and Information, Radboud University, 6500 HE Nijmegen, P.O. Box 9104, The Netherlands*
[b] *Department of Mathematics for Science and Technology, University of Minho, 4800-058 Guimaraes, Portugal*

Abstract

Many of our daily activities are supported by behavioural goals that guide the selection of actions, which allow us to reach these goals effectively. Goals are considered to be important for action observation since they allow the observer to copy the goal of the action without the need to use the exact same means. The importance of being able to use different action means becomes evident when the observer and observed actor have different bodies (robots and humans) or bodily measurements (parents and children), or when the environments of actor and observer differ substantially (when an obstacle is present or absent in either environment). A selective focus on the action goals instead of the action means furthermore circumvents the need to consider the vantage point of the actor, which is consistent with recent findings that people prefer to represent the actions of others from their own individual perspective. In this paper, we use a computational approach to investigate how knowledge about action goals and means are used in action observation. We hypothesise that in action observation human agents are primarily interested in identifying the goals of the observed actor's behaviour. Behavioural cues (e.g. the way an object is grasped) may help to disambiguate the goal of the actor (e.g. whether a cup is grasped for drinking or handing it over). Recent advances in cognitive neuroscience are cited in support of the model's architecture.

1. Introduction

Many of our activities in daily life involve some form of cooperation with another person, for example, carrying an object together or holding a cup so that the other can pour in a drink. These joint actions not only require coordinating each others actions at the motor control level (e.g. Burstedt, Edin, & Johansson, 1997), but also incorporating the actions of others in the planning of one's own actions. To do this, the other's actions must first be recognised. Central to the issue of action observation is the question of how an observed action can be linked to the observer's own actions. There are two main theories of how this could work in humans. According to 'theory of mind' humans are able to recognise another's actions because they possess a common sense model of human behaviour (Frith & Frith, 1999; Grèzes, Frith & Passingham, 2004). According to 'simulation theory' the observer simulates what the other is doing by using her/his own action system (Goldman, 1992).

The discovery of 'mirror neurons' showed that the action system is indeed involved in action observation (Gallese and Goldman, 1998). Mirror neurons have the interesting property of firing selectively both during observation and execution of goal-directed actions (Rizzolatti, Fogassi, & Gallese, 2001; Iacoboni, Moinar-Szakacs, Gallese, Buccino, Mazziotta & Rizzolatti, 2005). Brass, Bekkering, and Prinz (2001) have found that actions are initiated more quickly when observing a congruent action than an incongruent action. This finding also suggests that the action system is involved in action observation.

Action observation also plays a crucial role in imitation learning, which is a major topic in robotics (e.g. Breazeal, Buchsbaum, Gray, Gatenby, & Blumberg, 2005; Schaal, Peters, Nakanishi, & Ijspeert, 2004; for an overview see Schaal, 1999). A commonly used technique is a direct mapping of observed joint angles onto a set of learned action primitives in order to associate actions of the self to observed actions (Schaal, 2003). However, such an approach can only work when the bodies of the observed agent and the observer are similar. For humans this assumption is not very realistic not only because the relative lengths of the limbs may differ (adults

and children), but also because humans are capable of inferring action goals of highly dissimilar bodies. For example, when an octopus grasps an object with one of its arms (Yekutieli, Sagiv-Zohar, Aharonov, Engel, Hochner, & Flash, 2005; Yekutieli, Sagiv-Zohar, Hochner & Flash, 2005), observers are still able to infer the target of the movement.

Bekkering, Wohlschläger, and Gattis (2000) have argued that the goals of an action are more important in imitation—and in action observation in general—than the means with which the particular goals are being accomplished. Van Schie, Mars, Coles, and Bekkering (2004) found that observation of left and right hand movements of a facing actor activated motor areas contra-lateral to the side of the observed action (corresponding to the ipsi-lateral hand). Thus, the observer's brain activation reflects what the observer would have done if he/she had actively performed the action, instead of representing the task from the perspective of the actor. Neurophysiological evidence suggest that the intentions of observed actions are also important during action observation: the activity of mirror neurons of the macaque monkey (Fogassi, Ferrari Gesierich, Rozzi, Chersi, & Rizzolatti, 2005; Rizzolatti et al., 2001) and the human brain (Iacoboni et al., 2005) are modulated by the intentions associated to the observed actions. Inferring the goal of an action has substantial advantages over simply mapping the means of an action directly onto the observer's motor repertoire. First, goal inference can overcome large differences in bodily measures (Calinon, Guenter, & Billard, 2005). Second, it also allows the observer to use different means for achieving an action goal than the agent being observed (Erlhagen, Mukovskiy, Bicho, Panin, Kiss, Knoll, Van Shie, & Bekkering, in press).

Recently, Oztop, Wolpert, and Kawato (2004) have proposed a computational model of action observation. In their model the observed movement trajectory is extrapolated into the future using forward modelling of a hypothesised action. The hypothesised action that yields the best prediction is interpreted as equivalent to the observed action. It is assumed that forward modelling can be used to predict the movements of another person, which is very different from predicting proprioceptive feedback (afferent signal) of self-produced movements. This is problematic for several reasons. For one, it is not clear how such an approach is able to deal with the different vantage points of observer and actor. Another key issue is the amount of time that may be anticipated. In joint action it is very useful to know what the other is doing ahead of time in order to respond adequately. Many studies have shown that humans are indeed capable of anticipating another's behaviour (e.g. Flanagan & Johansson, 2003; Knoblich & Jordan, 2003; Wilson & Knoblich, 2005). When co-ordinating each other's actions the time spans involved can be quite large (a single reaching movement typically lasts between 0.5–2.0 s and manipulation of several objects lasts even longer). Forward modelling may work well to get rid of internal delays (Mehta & Schaal, 2002) but it is unlikely that large time spans can be reliably covered that way, which leads to noisy results (see Figs. 5 and 6 in Oztop et al., 2004). Critically, these problems may be circumvented by inferring the goal of an action rather than the action itself: action goals are not view-point dependent and they cover sufficiently large time spans because an action goal is the end result of an action.

Unfortunately, action goals are not directly observable. In most situations accurate inferences regarding action goals require additional knowledge. For example, when observing another person, the movement kinematics only contain information about the intended location. To infer the goal of an observed movement it is necessary to use information about the potential goals, which the observer may have acquired earlier.

In this paper, we present a computational model of how the goal of an action may be inferred from action observation. We focus on how knowledge about the task, the environment and the observer's action repertoire are being used to make these inferences. To do so we use a Bayesian framework because it allows us to separate prior knowledge from the information that is directly observable. In the process we indicate how the observer's personal preferences can disambiguate the goal inference at the cost of introducing a bias to the observer's preferred action goal.

2. Model architecture: the construction task

We illustrate the model architecture by means of a construction task, where a model has to be built from a collection of elementary components consisting of nuts, bolts, slats, etc. (Fig. 1). In order to perform such a task, knowledge is required about what one can do with the components. This includes the specific actions that must be accomplished, as well
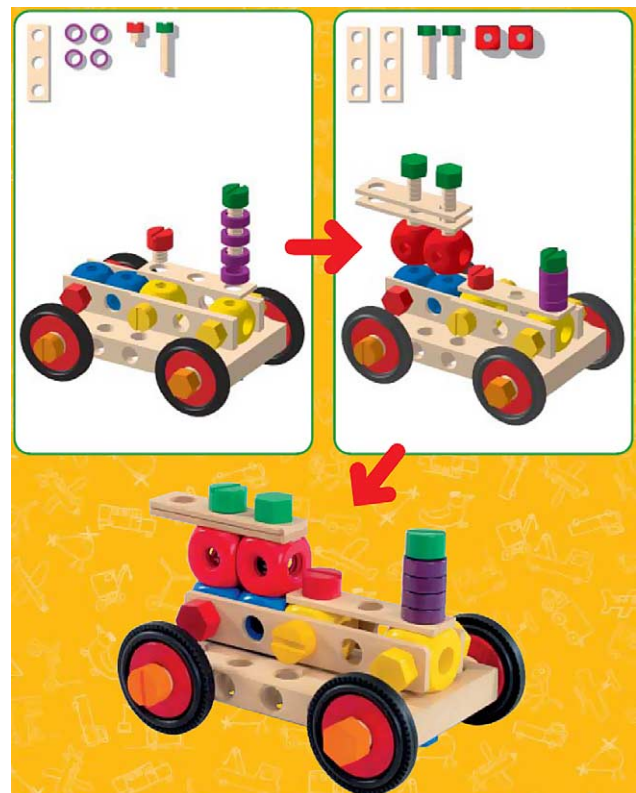

Fig. 1. Part of a building plan in a construction task.

as the manner in which they may be performed. Simply put, the *what* and the *how*. The knowledge about how acts need to be performed and what they accomplish is, in principle, available from prior experiences (we do not consider learning). In particular, there is some action repertoire with which the components may be manipulated. For each of those actions, it is known which components are needed and what the result of performing that action is. On the other hand, the knowledge about what needs to be done is not present a priori and is obtained from studying the building plan (Fig. 1). In the most extreme case, the building plan is nothing more than a picture of the completed model and the actions to undertake must somehow be deduced from this picture. Because of the large number of combinations with which the components may be assembled, the person building the model needs to make a number of choices without being certain that they will lead to the desired final result. This uncertainty arises from the fact that one can only reliably 'predict' the result of a few construction steps. Thus, one must evaluate, in some way, which of all possibilities most likely leads to the desired final construction. Our conjecture is that such evaluations are coarse or even ambiguous and that any choice is biased by the preferences of the person building the model.

When co-operatively building a model, the actions of the co-actor should be taken into account because they influence what the best course of action is. We assume that inferences about the co-actor's actions take place at the action goal level. This allows anticipation of the co-actor's behaviour over much larger time spans. However, inferring an action goal is only possible if the observations are combined with knowledge about the task, because the observations do not contain information about the action goals themselves. This still leaves open the questions what an action goal is, and how it is represented in the human brain. Although these questions are beyond the scope of the current paper, any representation of action goals should be independent of the means with which the action goals are being achieved. In other words, it is the end result that counts. We use a transition between states of the world $s_t = i \rightarrow s_{t+1} = j$ to represent action goals. For example, in the construction task the action goal is completely determined by the initial state (the unassembled components) and the state after performing the corresponding action (the assembled set of components). The advantage of this representation is that it does not depend on the vantage point of the observer, so, ideally, it is equivalent for each observer. The states of the world can never be represented internally in their full complexity. Fortunately, most of their parameters do not change or they change in a way that is irrelevant for a given task. In addition, many relevant parameters do not need a sustained internal representation but are available through observation. For example, the locations of objects may be observed directly and they do not need to be memorised with high accuracy.

In the construction task the number of components is limited and so is the number of ways in which they may be combined. Therefore, the number of action goals is also limited and we can arbitrarily assign a number i to them. This does not mean

that the number of actions, which are performed to accomplish these goals, are limited. However, human movements tend to be stereotypical for any given task in the sense that the movement kinematics are similar (Desmurget, Pélisson, Rossetti, & Prablanc, 1998) and that the employed end postures are similar despite the large number of degrees of freedom of the motor system (Gréa, Desmurget, & Prablanc, 2000). Also, the variability in end posture is small compared to the variability during the movement (Harris & Wolpert, 1998). Such observations have been explained by assuming that movements are generated using a limited set of primitives (Thoroughman & Shadmehr, 2000; Rosenbaum, Meulenbroek, Vaughan, & Jansen, 2001). This is also supported by the fact that canonical neurons and mirror neurons in the premotor areas of the macaque monkey brain (Fogassi et al., 2005) and human brain (Gangitano, Mottaghy, & Pascual-Leone, 2004) are selectively sensitive to particular action sequences. Certainly, this assumption decreases the computational load considerably making it popular in robotics (e.g. Paine & Tani, 2004). Thus, it seems that humans are capable of only a limited number of actions at a given time and that we can distinguish them in terms of the action goals they achieve. In our formalism we define a set of action alternatives $\mathcal{A}_t = \{A_1, \ldots, A_K\}$ such that the result of each action is a new state of the world $(A_k : s_t \rightarrow s_{t+1})$ as illustrated in Fig. 2A. We use the term *action alternative* rather than *action primitive* because we think of each action alternative as being some composition of action primitives. The action alternative 'screw bolt c_1 in nut c_2'', for instance, involves several grasping, holding, orienting and twisting movements. It is conceivable that different action alternatives entail the same change of states. For example, an object may be grasped using a precision grip (using the tips of thumb and index finger only) or a full grip. We regard action alternatives as different when the means with which a certain action goal is being achieved differ. In the construction task the relevant dimensions of the state of the world are determined by the set of components $C_t = \{c_1, \ldots, c_N\}$, where each component is some aggregate of the elementary building blocks. Note that each action alternative represents a 'move' from one collection of components to another. As such, the action alternatives define a Markov process (Sutton & Barto, 1998). Suppose A_1 is
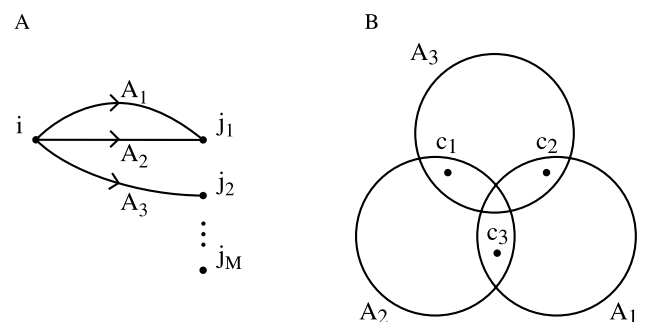

Fig. 2. (A) Every action alternative A_k corresponds to a unique transition $i \rightarrow j_m$, but the converse is not true. (B) Each action alternative A_k requires several components c_l indicated by the circles. On the other hand, each component c_l may be used for different action alternatives. For example, c_1 belongs to both A_2 and A_3.

the action alternative 'screw bolt c_1 in nut c_2', then A_1 transforms $\{c_1, c_2, ..., c_N\} \rightarrow \{<c_1, c_2>, c_3, ...c_N\}$ where $<c_1, c_2>$ denotes the newly assembled piece. Clearly, the new piece can only be assembled if c_1 is a bolt and c_2 is a nut, otherwise the assembly fails and the 'new' state of the world is identical to the previous one. We denote the subset of components on which the action alternative A_k is applicable by C_k (Fig. 2B).

2.1. Action observation

When observing an elephant grasping a peanut with its trunk, it is clear that the means used to grasp the peanut cannot be directly represented by the action repertoire of a human observer. If we do not wish to assume that the observer has knowledge about the action repertoire of an observed actor, the best an observer can do is to anticipate the action goal itself. However, the action goals are not directly observable. For example, when the goal is 'screw the red bolt in the blue nut', the observed motor act could be a reaching movement towards the red bolt (Fig. 3). Observations of the movement kinematics contain information about the target well before it is reached (Cuijpers, Smeets, & Brenner, 2004), so that the observer may anticipate the target of the movement. Whether the inferred target is the actual target of the movement is uncertain because of perceptual uncertainties and distractors in a cluttered environment. Many studies have shown that humans take into account these uncertainties and often combine them in a

statistically optimal way (Ernst & Bülthoff, 2004; Trommershäuser, Gepshtein, Maloney, Landy, & Banks, 2005). This can be modelled by defining the likelihood that the 'evidence' from observations occurs given the hypothesised target of the movement (denoted by $p(o_t|c_n)$ in Fig. 3). In order to obtain the target likelihood, Oztop et al. (2004) have proposed to use forward modelling of the hypothesised action to predict the perceptual consequences of future states of the motor system. Only the correct hypothesis will yield good agreement with subsequent observations and, thus, single out the correct action alternative. If knowledge about the task is available, the future states are known in advance because the potential targets are known. This can be used to define the target likelihood directly. Consequently, the target likelihood depends on some distance measure between the location (and velocity) of the actor's effector and those of the potential targets.

The target likelihood still does not uniquely determine the action goal. However, if the observer knows which targets are manipulated by each of her/his own action alternatives, it is possible to associate the target likelihood to the observer's action alternatives (Fig. 2B). Note that the action alternatives, which are associated to the target likelihood in this way, may represent means that are completely different from the observed movement. Observations may provide information about the target of a movement, but not about what the target object will be used for. Thus, in order to infer the action goal prior knowledge is required, either about the person performing the task ('she always combines red with blue objects') or about the desired final state. In joint action, the latter is known because it is common to all participants cooperating in a joint task. We model these personal interpretations by introducing a probability distribution $p(i \rightarrow j)$ over the action goals $i \rightarrow j$. This distribution is the result of action planning in an earlier stage (discussed below). Fig. 3 illustrates an example of action observation in the construction task. Both the red bolt and the green nut assign a high likelihood to the 'evidence' from observation as they are both in the direction of movement. But if the observer has knowledge about which action goals lead to the desired final state, the correct action goal may be inferred.

2.2. Planning action goals

When performing a task, typically a sequence of action alternatives need to be executed in order to reach the desired final state. For example, when making coffee, the filter must be placed in the coffee maker before the coffee itself. Typically, two types of errors have been observed in neurological patients: sequence errors (the coffee is put in the coffee maker before the filter) and substitution errors (sugar is put in the filter instead of coffee). In the construction task similar mistakes may occur but the situation is more complicated, because the number of ways in which the components may be combined can be much larger. Moreover, it is not immediately clear which of all possibilities will lead to the desired final state—a completed model. Clearly, a brute force search through the entire state space is not the solution, because of the combinatorial explosion of possibilities and because it is
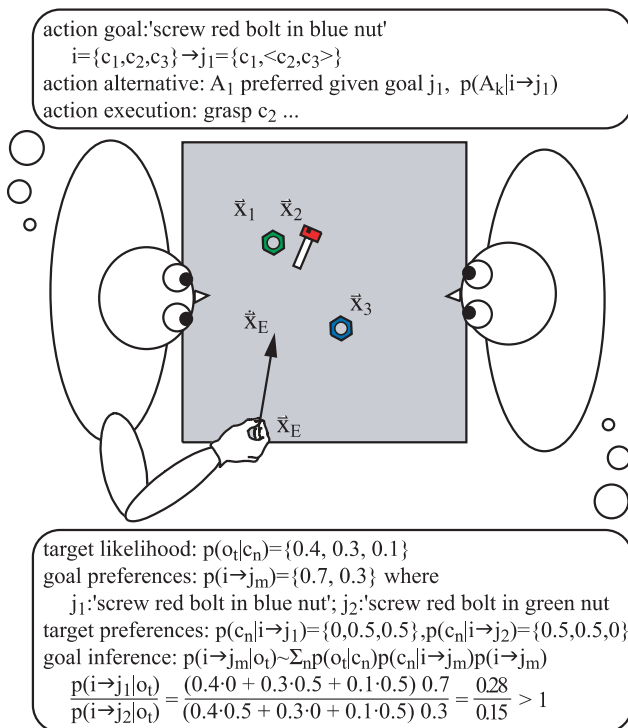

action goal:'screw red bolt in blue nut'
   $i=\{c_1,c_2,c_3\}\rightarrow j_1=\{c_1,<c_2,c_3>\}$
action alternative: A_1 preferred given goal j_1, $p(A_k|i\rightarrow j_1)$
action execution: grasp c_2 ...

target likelihood: $p(o_t|c_n)=\{0.4, 0.3, 0.1\}$
goal preferences: $p(i\rightarrow j_m)=\{0.7, 0.3\}$ where
  j_1:'screw red bolt in blue nut'; j_2:'screw red bolt in green nut
target preferences: $p(c_n|i\rightarrow j_1)=\{0,0.5,0.5\}$,$p(c_n|i\rightarrow j_2)=\{0.5,0.5,0\}$
goal inference: $p(i\rightarrow j_m|o_t)\sim\Sigma_n p(o_t|c_n)p(c_n|i\rightarrow j_m)p(i\rightarrow j_m)$
$$\frac{p(i\rightarrow j_1|o_t)}{p(i\rightarrow j_2|o_t)} = \frac{(0.4\cdot 0 + 0.3\cdot 0.5 + 0.1\cdot 0.5)}{(0.4\cdot 0.5 + 0.3\cdot 0 + 0.1\cdot 0.5)}\frac{0.7}{0.3} = \frac{0.28}{0.15} > 1$$

Fig. 3. Action planning and action observation. The actor (left side) has decided to screw the red bolt (c_2) in the blue nut (c_3) and reaches for the red bolt. The observer (right side) observes the locations of hand ($\vec{x}_E$) and targets ($\vec{x}_i$), and the speed of the actor's hand ($\dot{\vec{x}}_E$). The action goal is inferred using prior knowledge about the potential goals $p(i\rightarrow j_m)$ and which components are involved in each action goal $p(c_n|i\rightarrow j_m)$.

unlikely that the entire state space is known. There are plenty of examples such as making a jig saw puzzle of 1000 pieces. Therefore, we use an evaluation function to model the knowledge about which action goals lead to the desired final state (also called heuristic or utility function). The evaluation function defines an ordering in the state space. This ordering circumvents the need to search the entire state space because it allows one to judge whether one state is 'in the direction of' the desired final state or not. The evaluation value is not necessarily unique for each configuration of components: there are more ways to reach the final state. In that case, the personal preferences can disambiguate any decision. We model these personal preferences by a probability distribution $p(i \rightarrow j)$ over the action goals $i \rightarrow j$. By weighing the personal preferences with the evaluation function we obtain a distribution over the action goals conditional on the final state f.

2.3. Selecting action alternatives

Based on the distribution over the action goals a decision can be made about which action goal to choose. However, there may be more than one action alternative that accomplishes this goal. Again, additional knowledge is required to choose the best alternative. There may be both internal and external reasons to choose one action alternative over another. For example, if one wishes to drink from a cup it is convenient to grasp it by the handle (internal constraint) but only when the cup has a handle (external constraint). These constraints lead to graded preferences of the action alternatives, which we model by a probability distribution over the action alternatives $p(A_k|i)$. Of all action alternatives that lead to the desired final state, the preferred one is chosen.

3. Model details

3.1. Definitions

1. s_t: state of the world at time t that is identified with a unique number $i \in \mathbb{N}$.
2. o_t: vector of observables during action observation.
3. $\mathcal{C}_t = \{c_1, ..., c_N\}$: set of components, where each component c_n is either a building block or an aggregate of building blocks.
4. $\mathcal{A}_t = \{A_1, ..., A_K\}$: set of action alternatives at time t, where each action alternative involves a state transition $(A_k:s_t \rightarrow s_{t+1})$.
5. $C_k \subseteq \mathcal{C}_t$: subset of all available components that are associated to action alternative A_k.
6. $V^f(s_t)$: evaluation function associating a value to each state of the world s_t when the desired state is $s_\infty = f$.
7. $\kappa_{ij} = \{k|A_k(i) = j, A_k \in \mathcal{A}_t\}$ list of indices of the action alternatives A_k which result in the same state $s_{t+1} = j$ given the current state $s_t = i$.
8. $\lambda_k = \{l|c_l \in C_k, C_k \subseteq \mathcal{C}_t\}$ list of indices of components c_n that are associated to action alternative A_k.

9. r_{ij}^l: reward during action observation for inferring state $s_{t+1} = j$ when it turns out to be $s_{t+1} = l$.

3.2. Action observation

The action goal is represented by a state change from $s_t = i$ to $s_{t+1} = j$, in short denoted $i \rightarrow j$. To infer the action goal of an observed action we assign a probability distribution over the possible state changes for a given set of observations o_t and a desired final state f:

$$p(i \rightarrow j|f, o_t) = p(s_{t+1} = j|s_t = i, f, o_t)$$
$$\propto p(o_t|s_t = i, s_{t+1} = j, f)p(s_{t+1} = j|s_t = i, f) \quad (1)$$

$$p(i \rightarrow j|f, o_t) = p(o_t|i \rightarrow j)p(i \rightarrow j|f). \quad (2)$$

Here, we have used Bayes rule to invert the conditional probability. The first term on the right hand side of Eq. 1 denotes the likelihood that the observed data occurs for a given state change from $i \rightarrow j$. Formally, this term is conditional on the final state f, but since the observations usually only contain information about the state change from $i \rightarrow j$, we may assume that the observations are independent of f (Eq. (2)). The second term is the prior distribution representing the goal preferences of the observer. On the other hand, if the observer has an action plan of his own, the probability will be larger for those $i \rightarrow j$ conforming to the action plan that leads to the desired final state f. Therefore, this probability is conditional on the final state f.

During action observation (Fig. 4A) the state $s_{t+1} = j$ in Eq. (1) is the hypothesised result of the observed action. Since, the observed action may not exist in the observer's own action repertoire, we assign a likelihood distribution $p(o_t|A_k, i \rightarrow j)$ over the observer's own action alternatives that result in the same state $(A_k(i) = j)$. This distribution denotes the likelihood that the observations o_t occur when the observer would perform action A_k. The likelihood is high whenever the observer's own action alternatives match the observed action. The likelihood distribution represents the observed action. It is related to the likelihood distribution across action goals in Eq. (2) by:

$$p(o_t|i \rightarrow j) = \sum_{k \in \kappa_{ij}} p(o_t|A_k, i \rightarrow j)p(A_k|i \rightarrow j), \quad (3)$$

where $\kappa_{ij} = \{k|A_k(i) = j, A_k \in \mathcal{A}_t\}$. The second term on the right hand side of Eq. (3) denotes the prior distribution reflecting the observer's preferences for the various action alternatives.

In the construction task each action alternative applies to a subset of the components $\mathcal{C}_t$, denoted by the subset C_k. Let $\lambda_k = \{l|c_l \in C_k, C_k \subseteq \mathcal{C}_t\}$, then we can write:

$$p(o_t|A_k, i \rightarrow j) = \sum_{l \in \lambda_k} p(o_t|c_l, A_k, i \rightarrow j)p(c_l|A_k, i \rightarrow j). \quad (4)$$

The second term on the right hand side of Eq. (4) is the probability that component c_l is the target of the movement when the actor intends to reach state j using action alternative
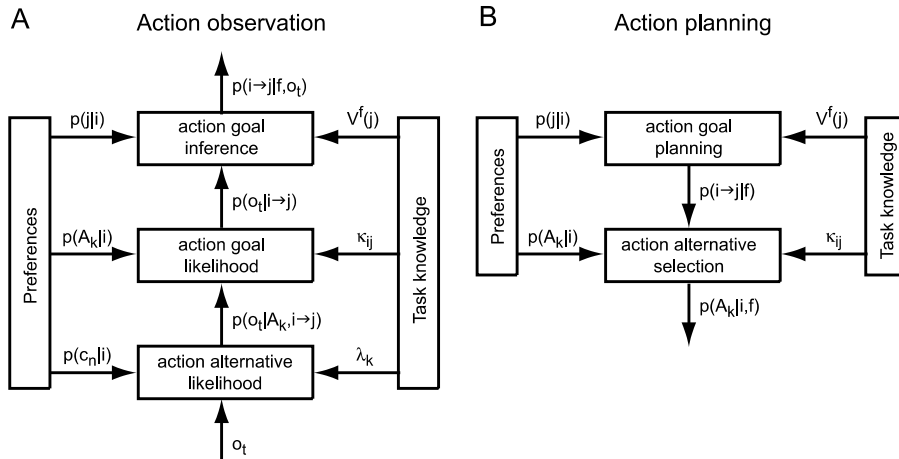
Fig. 4. Schematic representation of action observation (A) and action planning (B) indicating how task knowledge and personal preferences are combined.

A_k. This probability is proportional to the preference for that component but only if it is actually used by action alternative A_k (Fig. 2B):

$$p(c_l|A_k, i \rightarrow j) = \begin{cases} \dfrac{p(c_l|i)}{\sum_{n \in \lambda_k} p(c_n|i)} & \text{if} \quad l \in \lambda_k \wedge k \in \kappa_{ij} \\ 0 & \text{otherwise} \end{cases}. \quad (5)$$

The probability $p(c_l|i)$ indicates which component is preferred. These target preferences are conditional on the current state i because they could depend on the colour of the components, for instance, or on their distance from the actor's hand.

The first term on the right hand side of Eq. (4) is the likelihood that the observed movement is directed towards component c_l and that the means are consistent with action alternative A_k. This does not mean that the means of the observed action need to be in the observer's action repertoire, but the inference process will be better if they are. In order to find an expression for this target likelihood we need to specify what the observations o_t of the movement kinematics are. For simplicity we will only consider the locations and velocities of the effector ($\vec{x}_E(t)$) of the observed actor and those of the components ($\vec{x}_n(t)$). These observables carry information about which component will be manipulated but not about the means with which the component will be grasped. In that case it would be necessary to consider the locations and velocities of at least two effectors (e.g. the fingertips). We use the fact that the relative velocity between effector and target decreases with the distance from the target. Therefore, we introduce the following observable:

$$o_t = d_n(t) + \tau \dot{d}_n(t), \quad (6)$$

where τ is some constant and $d_n(t)$ denotes the distance between effector and component, defined by:

$$d_n(t) = |\vec{x}_n(t) - \vec{x}_E(t)| \quad (7)$$

We define the likelihood that the component c_n is the target of the movement as:

$$p(o_t|c_n, A_k, i \rightarrow j)$$

$$= \begin{cases} \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp -\dfrac{(d_n(t) + \tau \dot{d}_n(t))^2}{2\sigma^2} & \text{if} \quad l \in \lambda_k \wedge k \in \kappa_{ij} \\ 0 & \text{otherwise} \end{cases}, \quad (8)$$

where σ denotes the uncertainty. The likelihood is maximal when the effector is at the target after τ seconds. The constant τ can also be viewed as determining the relative contribution of the velocity information. Since the relative velocity decreases to zero as the effector approaches the actual target, the likelihood is high when the distance is decreasing. However, if the distance is decreasing too rapidly, the likelihood is small because it is likely that the effector will overshoot the target component. The form of Eq. (6) is rather arbitrary and many other definitions are possible. Clearly, the more knowledge about movement kinematics is incorporated, the better the target inference will be.

3.3. Making a decision

Thus far, we have only considered distributions, which assign likelihoods to either the action goals or the action alternatives. At some point a decision must be made in order to perform an action. A decision about which action goal to undertake may be achieved by maximising the reward that is obtained for inferring the correct action goal. We use the Bayes rule which maximises the expected reward with respect to the posterior distribution given by Eq. (2) when anticipating an action goal. We have:

$$j_{\text{anticipated}} = \arg \max_j \sum_l r_{ij}^l p(i \rightarrow j|f, o_t), \quad (9)$$

where r_{ij}^l is the reward for choosing j when l was the correct choice. A very simple reward function suitable for our purposes is:

$$r_{ij}^l = \begin{cases} 1 & \text{if } j = l \\ 0 & \text{if } j \neq l \end{cases}. \tag{10}$$

Then Eq. (9) reduces to:

$$j_{\text{anticipated}} = \arg \max_j p(i \rightarrow j | f, o_t). \tag{11}$$

Eq. (11) shows what the most likely action alternative is given some observations and a desired final state, but since observations change over time so does the planned action. The question then is when to make a decision. A decision may then be made when the likelihood $p(i \rightarrow j_{\text{anticipated}} | f, o_t)$ exceeds some predetermined level. If we take for this level a value proportional to the likelihood of the second best action goal $j_{\text{secondbest}}$, a decision is reached when

$$p(i \rightarrow j_{\text{anticipated}} | f, o_t) > \alpha p(i \rightarrow j_{\text{secondbest}} | f, o_t), \tag{12}$$

where α is some predetermined likelihood ratio.

3.4. Planning action goals

The planning of an action goal (Fig. 4B) without knowledge about the desired final state can only occur on the basis of some personal preferences. If the desired final state is known, the suitability of each action goal should be evaluated first. The personal preferences may still bias the results. We write for the second term on the right hand side of Eq. (1):

$$p(i \rightarrow j | f) \propto V^f(j) p(i \rightarrow j), \tag{13}$$

where $p(i \rightarrow j)$ are the action goal preferences and $V^f(j)$ is the evaluation function denoting the value of state $s_{t+1} = j$ when the ultimate goal is to reach final state f. For the construction task such an evaluation function could be the ratio between the number of correctly assembled components (not counting elementary components) to the square of the total number of components. This function is 0 when there are only elementary parts and 1 when all elementary parts are used in the completed model. This evaluation function does not associate unique values to each action goal because multiple alternatives may lead to the same action goal. In that case the personal preferences can bias the observer in favour of one action goal over another.

The second term of Eq. (13) indicates the action goal preferences. Since each action alternative uniquely determines the action goal $i \rightarrow j$, the action goal preferences are related to the action alternative preferences in the following way

$$p(i \rightarrow j) = p(s_{t+1} = j | s_t = i) = \sum_{k \in \kappa_{ij}} p(A_k | i), \tag{14}$$

where $\kappa_{ij} = \{k | A_k(i) = j, A_k \in \mathcal{A}_t\}$.

During action observation the anticipated goal of the observed actor may be used to plan a response. The observer simply plans an action goal using the anticipated state of the world instead of the current state:

$$p(s_{t+2} = j_{\text{response}} | s_t = i, f, o_t) = p(s_{t+2} = j_{\text{response}} | s_{t+1}$$

$$= j_{\text{anticipated}}, f). \tag{15}$$

The planned response depends only on $s_t = i$ because we assume that the probability of a state only depends on the previous state and not the state before that (Markov assumption). This simplification could be relaxed, if necessary. The right hand side of Eq. (15) is given by Eq. (13) with the current state i replaced by the anticipated state $j_{\text{anticipated}}$.

3.5. Planning action alternatives

Suppose the desired next state of the world is $s_{t+1} = j$ when the current state is $s_t = i$. The action alternatives that accomplish this task are those A_k for which $k \in k_{ij}$. The personal preferences $p(A_k | i)$ differentiate between these actions alternatives. The preference for an action alternative given the action goal $i \rightarrow j$ is:

$$p(A_k | i \rightarrow j) = \begin{cases} \dfrac{p(A_k | i)}{\sum_{l \in \kappa_{ij}} p(A_l | i)} & \text{if } k \in \kappa_{ij} \\ 0 & \text{otherwise} \end{cases}. \tag{16}$$

If there are no action alternatives that correspond to a transition from $i \rightarrow j$, the probability is zero.

In order to know what the desired action alternative is, we need to evaluate which of all possible states is most desirable. The desired state j has a likelihood given by Eq. (13), so that:

$$p(A_k | s_t = i, f) = p(A_k | s_t = i, s_{t+1} = j, f) p(s_{t+1} = j | s_t$$

$$= i, f) = p(A_k | i \rightarrow j) p(i \rightarrow j | f), \tag{17}$$

where $j = A_k(i)$ is the state that results from action alternative A_k. In Eq. (17) we have used that the probability of an action alternative A_k for a given transition from $i \rightarrow j$ is independent of the desired final state f.

4. Implementation for the construction task

So far, our model is phrased in rather abstract terms. This was done on purpose so that the model is flexible enough. For example, there are not many requirements on the action repertoire apart from the requirement that they cause transitions between states of the world s_t by changing some set of components c_t in some way. However, as soon as the action repertoire has been determined most other quantities are fixed as well. The only quantities that can still be freely chosen are the personal preferences $p(c_n | i)$, $p(A_k | i)$ and the evaluation function $V^f(s_t)$.

In order to implement our model to any particular task we need to define precisely what we mean by action alternatives, action goals and the states of the world. Any quantitative predictions will depend for a large part on these definitions. Moreover, the various model parameters need to be (experimentally) determined first. This is not the aim of the current

paper. Instead, we look for some more qualitative predictions that are experimentally testable. We do this by implementing our model to the following scenario: two people need to construct a model using Baufix building blocks. One person (the actor) builds the model and the other person (the observer) assists by pointing to (or handing over) the next, required building block as soon as possible. With this scenario we could measure the observer's reaction time and choices and compare the behaviour to our model.

4.1. Choosing the action repertoire

In order to implement our model to the construction task we must first define the observer's knowledge about the world and her/his action repertoire. As an example we use a 'world' with five elementary components: two bolts c_1 and c_2, two nuts c_3 and c_4 and a three-holed slat c_5 (see Fig. 6A). Next, we need to choose the action repertoire of the observer, because that will determine the action alternatives and the potential new states of the world (and therefore the action goals). We define three actions all of which combine two components into a single new component:

1. Put bolt X through hole of component Y
2. Screw nut X on bolt Y
3. Place nut X in front of hole of component Y

Clearly, the first action is only possible if the component X is a bolt and component Y has a hole through which the bolt can be put. Since there are two bolts, there are two possible alternatives to choose from. In addition, each bolt can be put in a hole from two directions and the component Y may posses more than one hole, so that there are 12 possible alternatives for action 1. Similar considerations apply to actions 2 and 3. If we wish to distinguish between all these new states of the world, we need to consider all action alternatives. For simplicity we fix the direction to 'from above' for action 1 ('from below' for action 3) and ignore the number of holes (always choose the first hole). Then we obtain eight different action alternatives A_k (two for action 1, four for action 2 and two for action 3). Each action alternative operates on a different pair of components $C^k \subseteq \mathcal{C}_t$.

The action alternatives determine the number of potential new states $s_{t+1} = j$ of the world and their associated action goals $i \rightarrow j$. In our simple example there are only three possible new states (j_1: 'bolt put through hole 1 of slat from above', j_2: 'bolt screwed in nut' and j_3: 'nut attached to hole 1 of slat from below'). If the desired final state f is known, the potential new states can be evaluated using the evaluation function $V^f(j)$. For our example, we define the final state f as a slat with equally directed bolts through the first two holes that are fixed in place with two nuts (similar to the cabin of the locomotive in Fig. 1).

We have captured the knowledge about the construction task and the possible actions that may be performed given the current state of the world, but we still need to set the personal preferences for the components ($p(c_n|i)$), the action alternatives ($p(A_k|i)$) and the action goals ($p(i \rightarrow j)$). The latter follow from

Eq. 14. This leaves 11 free parameters in our example (5 components and 8 action alternatives minus 2 because the probabilities have to add up to 1). The role of these preferences will be discussed later. For now we use the unbiased situation where all components and all action alternatives are equally preferred.

Now that both the knowledge about the construction task and the personal preferences are known the action goal may be inferred from an observed movement (see Fig. 4A). Eq. (8) gives the likelihood of the observations and action goal likelihood is inferred using Eqs. (2)–(5) and Eq. (13). If the ratio between the most likely and second most likely action goal exceeds a threshold α, a decision is made and a response can be initiated. Otherwise, a new observation is made and the process iterates. In Fig. 5 the entire procedure for inferring the action goal from an observed movement is summarised in the form of pseudo code.

4.2. Action goal inference as a function of movement time

In Fig. 6A a potential scenario is shown of how the action goal may be inferred during action observation. In this simulation the observed actor moves in a straight line towards component c_1. The movement velocity was taken from the tangential velocity of a real grasping movement (see Cuijpers et al., 2004). The movement velocity is bell-shaped function of time, so that the hand was almost stationary in the first and last 20% of the movement. The movement time is typically 1 s and is approximately independent of the distance of the target. During observation of the movement the observable o_t given by Eq. (6) changes as the movement unfolds. To determine an appropriate value for τ we plotted the rate of change of distance $\dot{d}$ as a function of the distance d for each target (Fig. 6B). In the final approach the relation between $\dot{d}$ and d is approximately linear for the correct component (blue line), so that the likelihood is maximal if the value of τ is equal to the negative of the slope ($\tau = 0.1$, black line). Using Eq. (8) and $\sigma = 0.2$, we obtain the component likelihood $p(o_t|c_n)$ for each component (Fig. 6C). In the first 20% of the movement when the hand is stationary the likelihoods are largest for components c_2 and c_4 because they are the nearest targets. From 20 to 40% of the movement the hand picks up speed and the likelihood is much larger for component c_2 than for any of the other components. From 40% of the movement and onward the correct target c_1 produces the largest likelihood. This shows that the component

```
Initialise 𝒞_t, action repertoire and f
Determine 𝒜_t and the subsets C_k
Determine action goals and initialise λ_k, κ_ij and V^f(j)
Initialise p(c_n|i), p(A_k|i) and calculate p(i → j)
Repeat
      Determine p(o_t|c_n, A_k, i → j) from observations
      Infer action goal likelihood p(i → j|f, o_t)
      if p(i → j_anticipated|f, o_t) > αp(i → j_secondbest|f, o_t)
            make a decision and initiate response
      endif
Until end of movement or response initiated
```

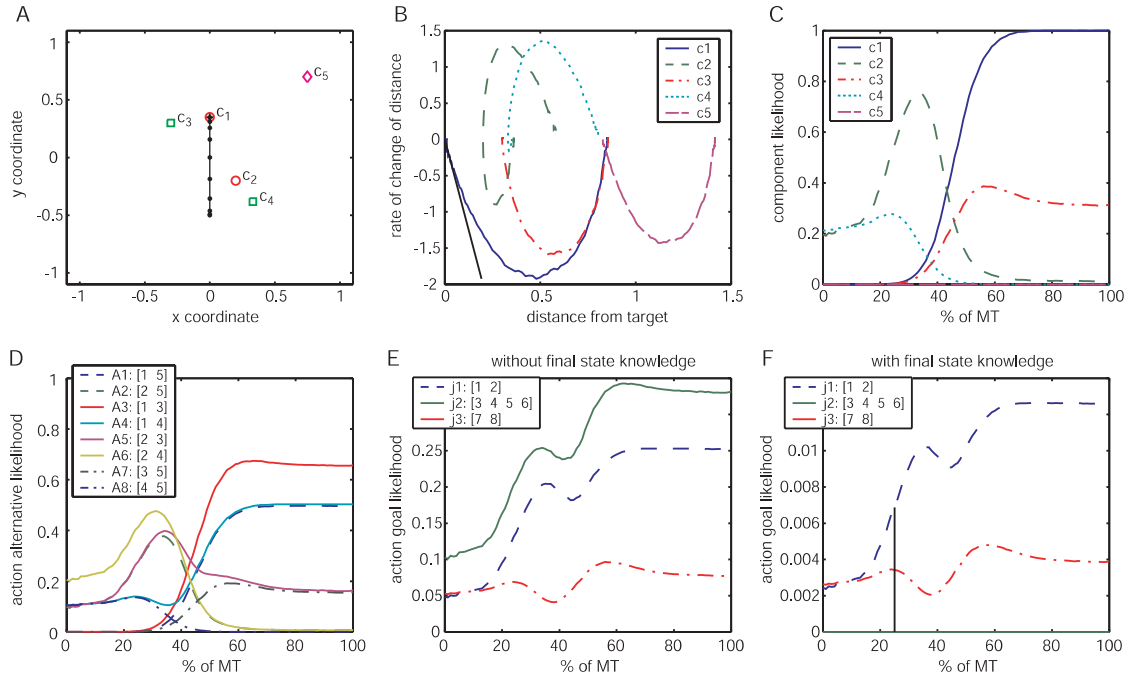Fig. 5. Pseudo code for action observation.

Fig. 6. Inferring the action goal during observation. *A*: scene layout of the bolts (circles), nuts (squares) and a 3-holed slat (diamond). The line indicates the movement trajectory with a dot at every 10% of the movement time. *B*: rate of change of distance plotted as a function of the distance of the hand from the target. The solid black line indicates the line $d + \tau \dot{d} = 0$, where $\tau = 0.1$. *C*: likelihood given each component as a function of time (in % of movement time). *D*: likelihood given each action alternative as a function of time (in % of movement time). The lists of components associated to each action alternative are indicated between brackets in the legend. *E*: likelihood given each action goal without using knowledge about the desired final state. The lists of action alternatives corresponding to each action goal are indicated between brackets in the legend. *F*: likelihood given each action goal using knowledge about the desired final state. The vertical line indicates the point in time where the likelihood ratio of the first and second largest likelihood exceeds the threshold $\alpha = 1.5$.

likelihood is unreliable for inferring the target of the movement in the first part of the movement. The same applies to action alternative likelihood given by Eq. 4 (Fig. 6D). The action alternative likelihood for A_k is large whenever the likelihood of one of the associated components C_k is large. In the first part of the movement the likelihood for action alternative A_6 is largest because it combines the most likely components c_2 and c_4. In the second part the likelihood of A_3 is largest because it combines the most likely components c_1 and c_3. Both these action alternatives have the same action goal $i \rightarrow j_2$ ('screw nut on bolt'), so that the likelihood for this action goal is largest throughout the entire movement (Fig. 6E). Thus, without knowledge about which action goal leads to the desired final state, action goal j_2 is the best choice. However, with knowledge about the desired final state (Eq. (2)) action goal j_2 is no longer an option because the bolt needs to be put through a hole of the slat first before screwing a nut on it. This is shown in Fig. 6F where only the correct action goals remain likely. The vertical, solid line indicates the point in time (about 25% of the movement time) where the likelihood ratio between the first and second best action alternative exceeds $\alpha = 1.5$ (Eq. (12)). Thus, the correct action goal has been inferred even before the correct target of the movement could be inferred (at about 40% of the movement time). The inferred action goal can be used to initiate an appropriate response. For example, the observer may assist the observed actor by handing over one of the missing components. Although the inferred action goal j_1 involves any of the components c_1, c_2 and c_5 (through A_1

and A_2), component c_5 is the only option because the observed movement is either directed to c_1 or to c_2. Because the action goal was inferred after about 25% of the movement time, there is still plenty of time for assisting.

4.3. The role of the personal preferences

To investigate the role of the personal preferences we created a scenario where the bolts c_1 and c_2 are placed symmetrically with respect to the movement that is headed for the slat c_5 (Fig. 7A). As a result the component likelihood is identical for the bolts c_1 and c_2 during the entire movement. The same applies to the action alternatives A_1 and A_2 and their associated action goal $i \rightarrow j_1$ as long as the observer is unbiased. In Fig. 7 we show what happens to the action alternative likelihoods (panel B) and the action goal likelihoods (panel C) at 60% of the movement time as the preference for c_1 increases relative to the preference for c_2. We keep the preferences for all other components unbiased ($p(c_n|i) = 1/5$ for $n = 3,4,5$). Since the probabilities add up to 1, we have that $p(c_1|i) + p(c_2|i) = pc_0 = 2/5$. It can be seen that the likelihood for A_1 (dashed line, top middle panel) decreases from almost 1 to about 0.3 as the preference $p(c_1|i)$ increases and vice versa for A_2. The likelihood for A_7, which does not involve components c_1 and c_2 remains unaffected. It may seem counterintuitive that the likelihood for action alternative A_1, which involves component c_1, is larger for small values of $p(c_1|i)$ than for large values. The reason is that when $p(c_1|i) = 0$ (and $p(c_2|i) = pc_0$), the
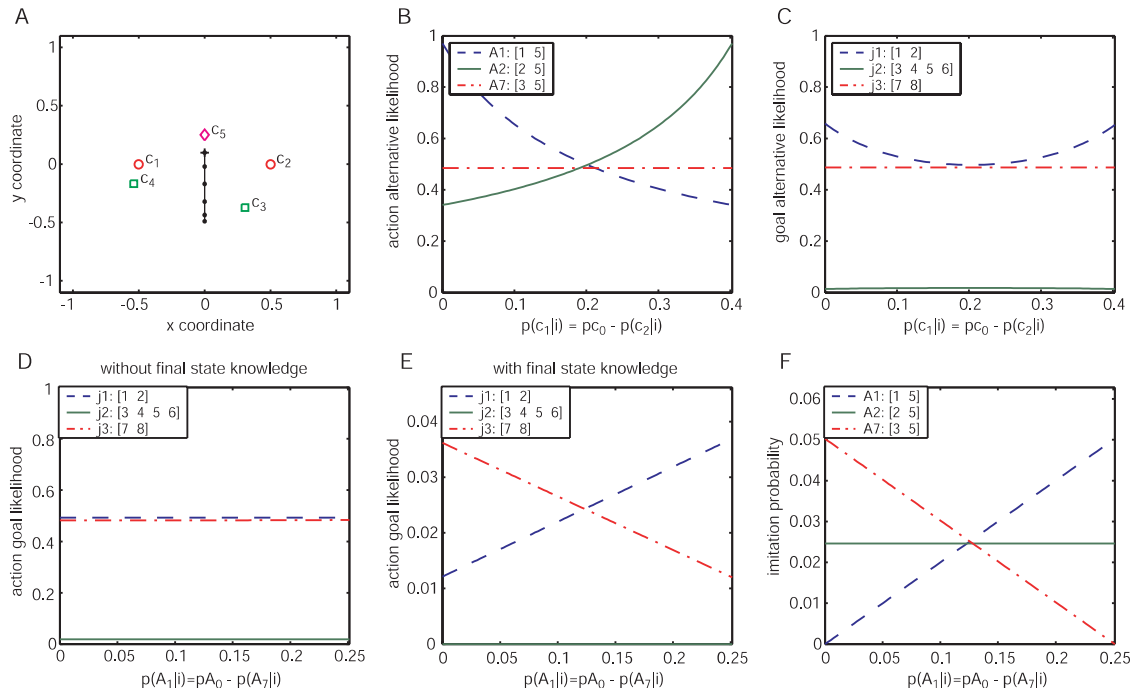
Fig. 7. Effect of changing the component preferences (B, C) and the action alternative preferences (D, E, F) on the action goal inference. *A*: scene layout of the bolts (circles), nuts (squares) and a 3-holed slat (diamond). The line indicates the movement trajectory towards c_5 with a dot at every 10% of the movement time. *B*: likelihood given each action alternative as a function of the preference $p(c_1|i)$ for component c_1 under the constraint that $p(c_1|i) + p(c_2|i) = pc_0 = 2/5$. The lists of components associated to each action alternative are indicated between brackets in the legend. *C*: likelihood given each action goal as a function of the preference for component c_1. The lists of action alternatives corresponding to each action goal are indicated between brackets in the legend. *D*: likelihood given each action goal as a function of the preference $p(A_1|i)$ for action alternative A_1 (without using knowledge about the desired final state f). The sum $p(A_1|i) + p(A_7|i) = pA_0 = 2/8$ is kept constant. *E*: same as D except that knowledge about the final state f is used. *F*: likelihood of the planned action alternative when the inferred action goal (panel D) is imitated.

likelihood for A_1 is completely determined by the likelihood that c_5 is the target of the movement, which happens to be the case. Similarly, the likelihood for A_2 is small because it is mostly determined by the likelihood that c_2 is the target of movement, which is not the case. Thus, changing the component preferences affects the weighting within each action alternative rather than between action alternatives. Since both action alternatives A_1 and A_2 are associated to the same action goal $i \rightarrow j_1$, its likelihood is larger in the case the component preferences are biased (Fig. 7C).

In a similar fashion, we can change the action alternative preferences $p(A_1|i)$ and $p(A_7|i)$ under the constraint that $p(A_1|i) + p(A_7|i) = pA_0 = 2/8$. As for the component preferences, changing the action alternative preferences will change the relative weights within the associated action goals: if $p(A_1|i) = 0$ (and $p(A_7|i) = pA_0$), the likelihood of the action goal $i \rightarrow j_1$ is determined by A_2. Since the most likely component c_5 is needed for both A_1 and A_2, the action goal likelihood is unaffected by our manipulation (Fig. 7D). However, the goal preferences defined by Eq. 14 are affected: $p(i \rightarrow j_1)$ increases linearly with $p(A_1|i)$. As a result, the action goal likelihood using knowledge about the final state f also increases linearly (Fig. 7E). The same applies to $p(i \rightarrow j_3)$ and $p(A_7|i)$. Thus, when the observations are ambiguous about what the action goal of the observed actor is, the personal preferences bias the action goal likelihood towards the preferred action goal. For example,

if the observer tries to imitate the action A_1, the response will be A_7 rather than A_1 when $p(A_1|i) < p(A_7|i)$ (Fig. 7F). Even if the inferred action goal is always correct ($i \rightarrow j_1$), the imitation response depends on the personal preferences: the observer imitates with A_2 rather than A_1 when $p(A_1|i) < p(A_7|i)$. Thus, whether or not an observer imitates an observed action with the same action depends on the personal preferences $p(A_k|i)$.

The personal preferences also affect the reaction time. If the observer sees a movement that conforms to her/his own preferences, the response is faster. We calculated the reaction times for the situation depicted in Fig. 6A, while varying the relative preferences for components c_1 and c_4. This is shown in Fig. 8 where the reaction time is plotted as a function of $p(c_1|i)$ under the constraint that $p(c_1|i) + p(c_4|i) = pc_0 = 2/5$.

5. Discussion and conclusions

In our model we have described how task knowledge and personal preferences can be combined to infer the action goal of an observed action. Although the action system is being used to interpret observed actions, our approach differs from simulation theory in two important ways. First, the action means are not being simulated and therefore they may differ from the observed action. Second, the goal inference is done from the observer's perspective and not from the vantage point of the observed actor. The observer uses her/his personal
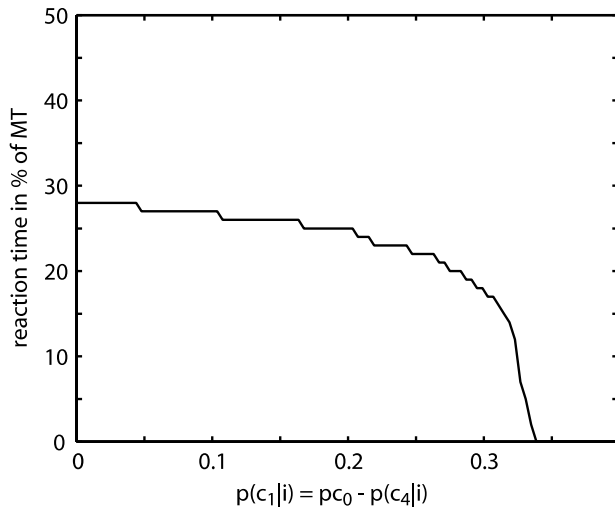
Fig. 8. Reaction time (in % of movement time) as a function of the preference $p(c_1|i)$ for component c_1 relative to the preference $p(c_4|i)$ of c_4. The sum of these probabilities is kept constant at $pc_0 = 2/5$.

preferences to infer the action goal, which will lead to mistakes if they differ from those of the observed agent. Although the inference process is from a personal perspective, the observations themselves are not. We used a viewpoint invariant measure to infer the target location, so that it does not matter whether the observed action is self-generated or not. The action goals may still be represented differently by each person because these representations are never compared directly but only through observations of the external world. In our model only action goals that are known to the *observer* can be inferred during action observation. For example, if the observer does not know that a cup can be used for drinking, the corresponding action goal preference is zero, so that he or she will never infer that the action goal is to drink from the cup. In Eq. 14 the action alternative preferences completely determine the action goal preferences. This expresses the situation where action goal preferences have been learned by exploring the world using one's own action repertoire. If this were the only way in which action goal preferences could be learned, the observer would be unable to infer an action goal that the observer cannot achieve her/himself. It is conceivable that action goals may be learned independently from the means with which these action goals can be achieved. In that case Eq. 14 needs to be modified to include knowledge obtained from other sources of information than the action system.

We used a one-dimensional distance measure as our metric for the observations. However, in general the observations will have more than one dimension. For example, when grasping an object with index finger and thumb, movement direction and hand pre-shaping variables could be estimated independently (Cuijpers et al., 2004; Desmurget et al., 1998). For each of these measures a goal-likelihood may be estimated, but how are they combined? We suggest a modular approach similar to the MOSAIC model (Wolpert & Kawato, 1998; Wolpert, Doya, & Kawato, 2003): the likelihoods for each independent measure are weighted and averaged according to their

responsibility. In the MOSAIC model the responsibilities indicate the relative importance of each measure for a given action alternative. This allows a flexible recruitment of the multi-dimensional observations for each action alternative. Triesch, Ballard, Hayhoe and Sullivan (2003) found that task relevant information is only processed when it is necessary. Therefore, the responsibilities may change over time so that they not only determine *what* sources of information are necessary but also *when* each source is needed.

In our model, task knowledge and personal preferences are both internal parameters that are being used to interpret the external sources of information. However, they are very different entities in terms of how they change over time. Task knowledge is more related to semantic knowledge in the sense that it represents knowledge about how objects can be manipulated. Although this knowledge can be learned, this kind of learning is a long term process, which we consider fixed for the duration of a given task. Personal preferences are, on the other hand, conditional on the current state of the world. Thus, they evolve over time. For example, if nearer components are preferred over more distant components, then changing their locations will also change the preferences. This opens up the possibility of changing another person's preferences and, thus, the possibility of influencing the other's action goal planning. For example, a component that is out of reach of another person may be placed in her/his vicinity so that the likelihood that this component will be used will increase. The personal preferences may also change over time due to learning. If the anticipated target of an action during action observation proves to be incorrect, the target preferences need to be updated accordingly, so that subsequent inferences are more likely to be correct. An interesting consequence is that observing another person will change the personal preferences over time to resemble those of the observed person.

We simulated the various likelihoods and probabilities in a scenario where one person builds a Baufix model and the other assists by handing over required objects. The simulations confirm that it is possible to infer the action goal even when the target of the movement and the associated action goal are ambiguous. In that case the model predicts that the response can occur before it is certain what the movement target is. If the observer knows the desired final state, she/he will make less errors and infer the correct action goal even if the 'evidence' from observations favours another solution. Changing the preferences of the components and action goals has a large effect on the response, but the effect on the inferred action goal is relatively small, especially if the observer does not know the desired final state. Finally, we showed that the reaction time will become less when the observed actor performs a movement that conforms to the personal preferences of the observer.

We devised our model of action observation by means of a construction task, but it is also possible to apply it as a framework to other tasks. For example, if we wished to model sign language, we could replace the action goals by hand signs, the action alternatives by the means to realise them and the 'component(s)' would be the person(s) to whom the sign

language was directed. The model framework then indicates the interactions between task knowledge, personal preferences and observations during action observation.

Acknowledgements

References

Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology*, *53A*, 53–164.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*, 3–22.

Breazeal, C., Buchsbaum, D., Gray, J., Gatenby, D., & Blumberg, B. (2005). Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots. *Artificial Life*, *11*, 31–62.

Burstedt, M. K. O., Edin, B. B., & Johansson, R. S. (1997). Coordination of fingertip forces during human manipulation can emerge from independent neural networks controlling each engaged digit. *Experimental Brain Research*, *117*, 67–79.

Calinon, S., Guenter, F., & Billard A. (2005). Goal-directed imitation in a humanoid robot. In Proceedings of the 2005 IEEE international conference on robotics and automation, Barcelona, Spain, April 2005.

Cuijpers, R. H., Smeets, J. B. J., & Brenner, E. (2004). On the relation between object shape and grasping kinematics. *Journal of Neurophysiology*, *91*, 2598–2606.

Desmurget, M., Pélisson, D., Rossetti, Y., & Prablanc, C. (1998). From eye to hand: Planning goal-directed movements. *Neuroscience and Biobehavioral Reviews*, *22*(6), 761–788.

Erlhagen, W., Mukovskiy, A., Bicho, E., Panin, G., Kiss, C., & Knoll, A., Van Schie, H., & Bekkering (in press). Action understanding and imitation learning in a robot-human task. In Proceedings of the international conference on artificial neural networks (ICANN 2005), Warsaw, Poland, September 2005.

Ernst, M. O., & Bülthoff, H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162–169.

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*, 769–771.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science*, *308*, 662–667.

Frith, C. D., & Frith, U. (1999). Interacting minds—A biological basis. *Science*, *286*, 1692–1695.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*(12), 493–501.

Gangitano, M., Mottaghy, F. M., & Pascual-Leone, A. (2004). Modulation of premotor mirror neuron activity during observation of unpredictable grasping movements. *European Journal of Neuroscience*, *20*, 2193–2202.

Goldman, A. (1992). In defense of simulation theory. *Mind and Language*, *7*, 104–119.

Gréa, H., Desmurget, M., & Prablanc, C. (2000). Postural invariance in three-dimensional reaching and grasping movements. *Experimental Brain Research*, *134*, 155–162.

Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *NeuroImage*, *21*, 744–750.

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, *394*, 780–784.

Iacoboni, M., Moinar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, *3*(3), e79.

Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29*, 1006–1016.

Mehta, B., & Schaal, S. (2002). Forward models in visuomotor control. *Journal of Neurophysiology*, *88*, 942–953.

Paine, R. W., & Tani, J. (2004). Motor primitive and sequence self-organization in a hierarchical recurrent neural network. *Neural Networks*, *17*, 1291–1309.

Oztop, E., Wolpert, D., & Kawato, M. (2004). Mental state inference using visual control parameters. *Cognitive Brain Research*, *22*, 129–151.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Neuroscience*, *2*, 661–670.

Rosenbaum, D. A., Meulenbroek, R. J., Vaughan, J., & Jansen, C. (2001). Posture-based motion planning: Applications to grasping. *Psychological Review*, *108*(4), 709–734.

Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*, *3*(6), 233–242.

Schaal, S. (2003). Dynamic movement primitives—a framework for motor control in humans and humanoid robotics. 2nd International symposium on adaptive motion of animals and machines (AMAM 2003).

Schaal, S., Peters, J., Nakanishi, J., & IJspeert, A. (2004). Learning movement primitives. In International symposium on robotics research (ISRR 2003), springer tracts in advanced robotics. Ciena, Italy: Springer.

Schie, H. T. van, Mars, R. B., Coles, M. G. H., & Bekkering, H. (2004). Modulation of activity in the medial frontal and motor cortices during error observation. *Nature Neuroscience*, *7*(5), 549–554.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. London: MIT Press.

Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, *407*, 742–747.

Triesch, J., Ballard, D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision*, *3*, 86–94.

Trommershäuser, J., Gepshtein, S., Maloney, L. T., Landy, M. S., & Banks, M. S. (2005). Optimal compensation for changes in task-relevant movement variability. *Journal of Neuroscience*, *25*(31), 7169–7178.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*(3), 460–473.

Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*, 1317–1329.

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

Yeutieli, Y., Sagiv-Zohar, R., Aharonov, R., Engel, Y., Hochner, B., & Flash, T. (2005). Dynamic model of the octopus arm. I. Biomechanics of the octopus reaching movement. *Journal of Neurophysiology*, *94*, 1443–1458.

Yeutieli, Y., Sagiv-Zohar, R., Hochner, B., & Flash, T. (2005). Dynamic model of the octopus arm. II. Control of reaching movements. *Journal of Neurophysiology*, *94*, 1459–1468.