Motor Simulation via Coupled Internal Models Using Sequential Monte Carlo

Haris Dindo*, **Daniele Zambuto**
University of Palermo
Computer Science Engineering
Viale delle Scienze, Ed. 6, Palermo, Italy
{dindo,zambuto}@dinfo.unipa.it

Giovanni Pezzulo
Consiglio Nazionale delle Ricerche
ILC-CNR and ISTC-CNR
Via S. M. della Battaglia, 44, Roma, Italy
giovanni.pezzulo@istc.cnr.it

Abstract

We describe a generative Bayesian model for action understanding in which inverse-forward internal model pairs are considered 'hypotheses' of plausible action goals that are explored in parallel via an approximate inference mechanism based on sequential Monte Carlo methods. The reenactment of internal model pairs can be considered a form of *motor simulation*, which supports both perceptual prediction and action understanding at the goal level. However, this procedure is generally considered to be computationally inefficient. We present a model that dynamically reallocates computational resources to more accurate internal models depending on both the available prior information and the prediction error of the inverse-forward models, and which leads to successful action recognition. We present experimental results that test the robustness and efficiency of our model in real-world scenarios.

1 Introduction

An authoritative view in computational motor control is that to act in goal-directed manner the brain employs internal models [Wolpert *et al.*, 1995; Miall and Wolpert, 1996] which are fundamental for understanding a range of processes such as state estimation, prediction and context recognition. Internal models could explain the human's ability to generate appropriate movement patterns under uncertain environmental conditions. As an example, in order to pick up an object without knowing its dynamics, it has been proposed that the Central Nervous System (CNS) uses a modular approach in which multiple internal models coexist and are selected based on the movement context or state, so that each internal model becomes an *expert* of a given motor action and context [Wolpert and Kawato, 1998].

In computational motor control there are two types of internal models. *Forward models* model the causal (forward) relation between actions and their consequences, and can be used to predict how the motor system's state changes in response to a given motor command under different contexts. *Inverse models*, known as controllers in engineering, perform the opposite transformation determining the motor command required to achieve some desired goal state. When a motor command is generated, an efference copy can be used to predict the sensory consequences under many possible contexts, where contexts correspond to execution of different motor actions. Predictions can be compared with actual sensory feedback to assess what context is more plausible given evidence. Each predictor can, therefore, be considered as a *hypothesis tester* for the context it models: the smaller the error in prediction, the more likely the context. Moreover, each predictor is paired with a corresponding controller forming a predictor-controller pair and the sensory prediction error is used to weight the relative influence of the paired controllers. We can therefore describe a goal-directed action as: (a) a controller, or inverse model, which determines the appropriate motor command to reach a certain goal state, coupled with (b) a predictor, or forward model, which predicts the consequences of the action in a given context.

In addition to motor control, many authors advocate that coupled inverse and forward models can be reused for explaining a number of complex social phenomena, ranging from action prediction and understanding [Wolpert, 2003], imitation learning in robotics [Oztop *et al.*, 2006; Demiris and Khadhouri, 2006], language production and speech understanding [Perkell *et al.*, 1997], just to name a few.

In this paper we present a computational model for action understanding, or the process of how we understand actions performed by others and their goals while observing them. Our model is based on the idea that recognition and understanding automatically emerge as the process of reusing one's own internal models "in simulation". This idea is consistent with the simulation theory of mindreading that has been advanced in the cognitive neuroscience literature, according to which, in order to understand another's actions and their underlying goals, humans encode observed actions in terms of their own motor repertoire, or in other words they "put themselves into another's shoes" [Gallese and Goldman, 1998]. The encoding of observed actions in terms of one's own permits to go beyond the observed movements and provides a direct, motor understanding of the goals of the performer. The link between two apparently unrelated processes of action understanding and motor simulation is suggested by recent neuroscientific studies related to the discovery of mirror neurons in the F5 area of the macaque brain, which are active both dur-

*Corresponding author

ing (transitive) action and the perception of the same action performed by others [Rizzolatti and Craighero, 2004]. Mirror neurons have provided (prima facie) support to the simulative view of action understanding, both for what concerns the understanding of the immediate goals of observed actions and distal intentions confirmed by empirical findings [Fogassi *et al.*, 2005].

Here we offer a computational account of how a motor simulation account of action understanding might work mechanistically (see also [Kilner *et al.*, 2007; Wolpert, 2003]). By embedding the ability to produce actions and predict their outcome via inverse and forward models, action recognition in our model can be described as an *inverse inference* process in which the generative model is "inverted" so as to pass from the observation of movements to the inference of which of the observer's inverse model(s) could have generated the observations. Since inverse models have associated goals, recognition of inverse models entails recognition of the (more plausible) action goals. Put in simple terms, the same internal models that an organism uses for performing goal-directed action can be re-enacted in simulation and used for inferring others' actions, intention and goals (and possibly imitating them). Forward models, or predictors, can be used as simulators of the consequences of an action, and when paired with an inverse model, a controller, a degree of discrepancy between what I observe and what I do (or just "imagine" of doing) can be produced which helps finessing the initial hypotheses about the observed action goals (and which inverse model could have produced it). Note that it is not a one-step process, but can be as dynamic as the action performance itself.

1.1 Computational complexity of motor simulation

In principle, the modular approach based on internal models gracefully solves the problem of action recognition and motor control. However, from practical point of view, it is unlikely and impractical that all models are maintained in parallel for the entire period of recognition. For each application context, there might be hundreds or thousands of internal models (and their number could easily grow in ambiguous situations where the environment does not provide sufficient information for recognition) making the problem of action recognition intractable with scarce resources. We must also consider the inherent diversity during the execution of the same action among different individuals, and the diversity in action execution when performed by the same individual during different trials. The complexity associated with tracking a huge number of possible models for each scenario, and their inherent stochastic nature, has so far hindered the development of efficient analytical solutions for motor simulation and action understanding in most but the simplest settings.

However, casting the problem of action understanding in a Bayesian framework permits to adopt efficient techniques for *approximate* probabilistic inference under the constraint of limited resources. We will show how, by adopting a sequential Monte Carlo scheme for inference, the process of action understanding can be efficiently and elegantly solved in a coherent computational framework. In addition, our model integrates different sources of information (i.e. affordances, context and preferences) and treat them in a homogeneous way as *Bayesian priors* that bias the initial allocation of resources and modulate the dynamics of our system.

1.2 Aims and structure of the paper

In this paper we present a complete and quantitatively specified probabilistic model for understanding predictive and motor phenomena during action observation. Our model has three key characteristics. First, internal models used in motor control are reused for motor simulation and action understanding at the goal level in a sound Bayesian framework. Second, prior knowledge of possible goals, as well as other contextual elements, are used to bias the initial action recognition inference, which is successively finessed during observation. Third, we use approximate Bayesian inference (particle filtering) rather than exact inference to efficiently track several competing action hypotheses. In other words, approximate Bayesian inference is offered as a plausible mechanistic implementation of the idea of motor simulation as the reuse of multiple forward-inverse models, which makes it feasible in real-time and with limited resources. Importantly, the second and third elements distinguish our model from previous ones proposed in the literature that are based on coupled forward-inverse models (e.g., [Cuijpers *et al.*, 2006; Dearden and Demiris, 2005; Haruno *et al.*, 2001; Lopes and Santos-Victor, 2005; Wolpert, 2003]), and pave the way to the adoption of the idea of motor simulation via coupled internal models in many real-world tasks such as surveillance, human-robot cooperation, learning by imitation and computer games.

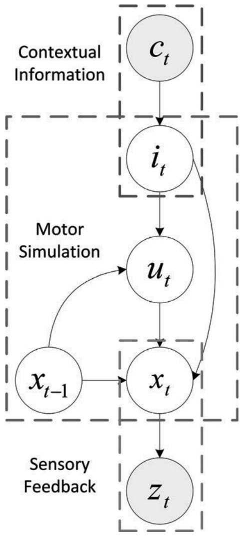
The paper is organized as follows. Sec. 2 describes the model in computational terms and discusses how the model deals with sudden changes during action performance and how it recognizes sequences of actions. Sec. 3 explains the computer simulations and human experiments that we have performed to assess the model, and it briefly discusses how it connects to the body of experimental literature in brain and cognitive sciences. Sec. 4 illustrates the conclusions and our ongoing work towards the modeling and recognition of distal intentions.

2 Computational model

Action understanding, when described as a motor simulation process, is influenced by three main factors: agent's repertoire of actions (represented as coupled forward-inverse models), contextual information, and observations of the movements of a demonstrator (i.e., agent performing the action). In motor simulation, it is the reenactement of one's own internal models, both inverse and forward, used for interaction that provides an understanding of what others are doing.

The entire process of action understanding via motor simulation can be cast into a Dynamic Bayesian Network (DBN) shown in Figure 1(a). DBNs are Bayesian networks representing temporal probability models where directed arrows depict assumptions of conditional (in)dependence between variables [Murphy, 2002]. As usual, shaded nodes represent observed variables while others are hidden and need to be estimated through the process of probabilistic inference.

In our representation, the process of action understanding



(a) Graphical model

c_t	context	$discrete \in \{1, \ldots, N_c\}$
i_t	goal-directed action index	$discrete \in \{1, \ldots, N_i\}$
u_t	control	*continuous*
x_t	state	*continuous*
z_t	observation	*continuous*

(b) Stochastic variables

$p(c_t)$	contextual information
$p(i_t \vert c_t)$	contextual induction
$p(u_t \vert x_{t-1}, i)$	inverse model
$p(x_t \vert x_{t-1}, u_t, i)$	forward model
$p(z_t \vert x_t)$	observation model (prediction error)

(c) Probability densities

Figure 1: Graphical model (DBN) for action understanding based on coupled forward-inverse models

is influenced by the following factors expressed as stochastic variables in the model (fig. 1b):

1. c: discrete context variable;
2. i: index of the agent's own repertoire of goal-directed actions: each action directly influences the activation of related forward and inverse models;
3. u: continuous control variable (e.g. forces, velocities, ...);
4. x: state (e.g. the position of the demonstrator's end-effector in an allocentric reference frame);
5. z: observation, a perceptual measurement related to the state (e.g. the perceived position of the demonstrator's end-effector on the retina).

Suppose we can extract the noisy measurements of the true state of the demonstrator, z_t, through some predefined perceptual process. The aim of the recognition process is to determine the goal-directed action, i_t, that the demonstrator is doing based on the observed state z_t. The action i is associated with a paired inverse-forward model, and it implicitly encodes the demonstrator's goal. The initial choice of which internal models to activate is biased by the a priori given contextual information, c_t. The context suggests actions (or sequences of actions) that may help explaining the observed behavior of the demonstrator. Distribution over possible contexts, $p(c_t)$, might be given through a predefined set of perceptual programs. Each action i_t is responsible of both generating a motor control u_t, given the (hidden) state x_{t-1} (inverse model), and predicting the next (hidden) state x_t, given the motor control u_t and the previous state x_{t-1} (forward model).

Figure 1c shows the conditional distributions which arise in the model. However, the semantics of the stochastic variables, and the concrete instatiation of the conditional distributions depends on the experimental setting, and we will provide an example in the section 3.

2.1 Probabilistic inference for action understanding

Let us denote with $\mathcal{X}_t$ the set of hidden variables at time t, and with $\mathcal{Z}_t$ the set of observed variables at the same time step. In general, we want to recursively estimate the posterior distribution $p(\mathcal{X}_t|\mathcal{Z}_{1:t})$ from the corresponding posterior one step earlier, $p(\mathcal{X}_{t-1}|\mathcal{Z}_{1:t-1})$. The usual Markovian assumptions lead to the following equation which, together with an a priori distribution $p(\mathcal{X}_0)$, provides the recursive formulation of the inference task [Murphy, 2002]:

$$p(\mathcal{X}_t|\mathcal{Z}_{1:t}) = \eta p(\mathcal{Z}_t|\mathcal{X}_t) \int p(\mathcal{X}_t|\mathcal{X}_{t-1}) p(\mathcal{X}_{t-1}|\mathcal{Z}_{1:t-1}) d\mathcal{X}_{t-1}$$

In our graphical model for action understanding (Fig. 1) the task is to recursively compute the posterior distribution over possible forward-inverse action pairs, $p(i_t|z_{1:t})$. This distribution can be efficiently obtained by marginalizing the posterior distribution over all hidden variables in the model through Monte-Carlo integration method. The following equations describe the observation and transition models, together with the a priori distribution over the set of hidden variables:

$$p(\mathcal{Z}_t|\mathcal{X}_t) = p(z_t|x_t) \quad (1)$$

$$p(\mathcal{X}_t|\mathcal{X}_{t-1}) = p(x_t|x_{t-1}, u_t, i) \cdot p(u_t|x_{t-1}, i) \quad (2)$$

$$p(\mathcal{X}_0) = p(x_0) \cdot p(c_0) \cdot p(i|c_0) \quad (3)$$

It is worth stressing how the coupled forward-inverse models naturally appear in the prediction model (equation 2 above).

However, in order to compute the most likely observed action, the recursive propagation of the posterior density $p(\mathcal{X}_t|\mathcal{Z}_{1:t})$ is only a theoretical possibility, and in general it cannot be determined analytically. We adopt *particle filters*, a Monte Carlo technique for sequential simulation, which allow to efficiently perform approximate computation of the posterior density with limited resources [Doucet *et al.*, 2000].

2.2 Particle filters

Functions that describe probability densities in real-world problems are typically nonlinear, and analytical solutions of the Bayesian inference problem are generally intractable. The key idea of particle filters is to represent the required posterior density function by a set of random samples with associated weights and to compute estimates based on these samples and weights. In our case, each particle represents a weighted hypothesis of an internal model activation in the action recognition task, and the weight of each particle is computed according to the divergence between the predicted state of the internal model the particle belongs to and the observed state; intuitively, severe discrepancies between predictions produced by coupled internal models and observed percepts will assign low weights to internal models less involved in explaining the current action observation.

Each random sample is therefore a distinct *hypothesis* that the agent tracks during the action recognition process. Let $\{x_t^k, u_t^k, i^k, c_t^k, w_t^k\}_{k=1}^{N_s}$ denote a random measure that characterizes our target posterior. The evidence provided by the perceptual process, z_t, is responsible of "correcting" the posterior distribution by integrating the observation model $p(z_t|x_t)$. Normalized importance weights w_t^k, recursively computed as the divergence between the predicted and observed state, together with the particle set, represent an approximation of the desired posterior probability. Intuitively, those parts of the hidden state in accordance with the current observation will be given higher weights and will thus exhibit peaks in the posterior distribution. Since those states have been produced by a goal-directed action, by marginalizing the final posterior distribution we obtain the required discrete distribution over motor primitives, $p(i_t|z_{1:t})$.

The resampling step in the classical particle filter, used to avoid the *particle impoverishment* problem in which the majority of particles' importance weights are close to zero, allows to focus computational effort on models providing plausible hypotheses (i.e. hypotheses in accordance with the observations) by pruning out less probable models and focusing computational resources to models that best explain the current observation through the activation of one's own internal models.

2.3 Novelty detection

A limitation of the model that we have presented is its poor performance when the demonstrator changes its action abruptly (e.g., a "feint" in a soccer game). The reason is that, in order to prevent the particle impoverishment problem, the resampling step of the particle filter usually assigns a high probability to a single model (or a small subset of models), and thus a whole set of particles will follow a unique dynamics. Although theoretically admissible, in many practical cases this behavior can induce misleading results. Consider for example the case when several actions have distinguishable features in the final stage of observation only. In such an ambiguous situation, the particle filter could start tracking a wrong dynamics. Even if the observations will at some point assign low weights to the predicted state, the system has no means to recover the true belief. A similar problem arises with switching dynamics (i.e. when the demonstrator rapidly changes the goal of its action).

This problem is intimately related to the process of *novelty detection* and it has a fundamental survival value since novelty often indicates dangerous or unexpected situations (and indeed the neural underpinnings of novelty detection have been widely investigated, see e.g. [Schultz *et al.*, 1997]). In principle, novelty can be detected when the observed world dynamics strongly differs from its predicted counterpart produced via active internal models. A potential solution is to explicitly model the transition between different models (i.e. $p(i_t|i_{t-1})$). At every time step, several particles will jump from one dynamics to another thus preventing the impoverishment by tracking a huge number of hypotheses. However, this random walk in the action space forces the recognition algorithm to process most of the actions at each time step, even those that have a low probability of being observed in a given context, thus making the algorithm computationally prohibitive and unsuitable to operate in real time with limited resources. A more sophisticated approach to the novelty detection problem is to populate the space with particles having different dynamics as soon as the particle impoverishment is detected. In this way, the computational burden of the overall algorithm is constrained and the particle filter can recover the true belief.

In order to detect a novelty, we use the informational theoretic measure based on *Kullback-Leibler (KL)* divergence [Bishop, 2006] between the current state belief, represented by the set of weighted particles, and the probability distribution induced by the current observation z_t. The algorithm will inject random particles from the state space when the KL divergence is larger than a given threshold.

We represent the current state belief distribution, $\mathcal{N}_x$, as a Gaussian with the first and the second moments (mean and variance) computed as below:

$$\mu_t^x = \sum_k w_t^k x_t^k \qquad \Sigma_t^x = \frac{1}{N}\sum_k w_t^k (x_t^k - \mu_t^x)(x_t^k - \mu_t^x)^T \tag{4}$$
$$N = 1 - \sum_k (w_t^k)^2$$

where w_t is the weight associated with each particle and N is a normalization factor. This expression reduces to an unweighted variance with the familiar $1/(N-1)$ factor when there are N identical non-zero weights.

In the same way, we summarize the present observation distribution ($z_t = \hat{x}_t$) as a Gaussian having the following statistics:

$$\mu_t^z = \hat{x}_t \qquad \Sigma_t^z = \sum_k (x_t^k - \mu_t^z)(x_t^k - \mu_t^z)^T \tag{5}$$

This distribution quantifies, through its covariance matrix, how agent's internal belief explains the current observation. In cases where the observation z_t directly provides a noisy measurements of the state $\hat{x}_t$ (as in the experiments we performed and described in the next section), this distribution can be directly compared to the estimated state in order to detect novelty by using KL divergence. In the general case, where observations and states are related through a complex non-linear transformation, the same technique can be easily applied in the domain of the observed variable z_t.

The degree of novelty is measured as KL divergence between these two distributions[1]:

$$D_{\mathrm{KL}}(\mathcal{N}_x \| \mathcal{N}_z) = \frac{1}{2}(\log_e(\frac{\det \Sigma_z}{\det \Sigma_x}) + \mathrm{tr}(\Sigma_z^{-1}\Sigma_x) + \\ + (\mu_z - \mu_x)^\top \Sigma_z^{-1}(\mu_z - \mu_x) - N) \quad (6)$$

When particle impoverishment is detected, the algorithm will inject a preset percentage of random particles by sampling i from $p(i|c_t)$ and x_t from $\mathcal{N}_x$. The injected particles, representing different system dynamics, will thus cover a subset of entire action repertoire conditioned on the current context c_t and present belief $\mathcal{N}_x$.

3 Experimental setup and results

To assess the real-time adequacy of our model, we compared its performance with human subjects in an action observation task, consisting in assessing what action the demonstrator is currently doing (i.e., approaching and grasping one among multiple target objects in the visual field, see fig. 2). The state x_t of the demonstrator is given by the 2D position of the end-effector (hand) relative to a fixed reference frame, and the observation z_t is provided by the noisy measurement of the 2D position by a low-cost motion capture device. Each intentional action is represented as a coupled forward-inverse model whose index is described through the stochastic discrete variable i_t. Inverse models, $p(u_t|x_{t-1}, i_t)$, are implemented as potential fields producing the control velocity vector, u_t, corrupted by a Gaussian noise with fixed variance, σ_i. In this formulation, each target object acts essentially as an attractor for the end-effector and the system automatically instantiates inverse models for *reaching* them. Forward models are based on a simple kinematics velocity model, $p(x_t|x_{t-1}, u_t, MP_t) = \mathcal{N}(x_{t-1} + \Delta t \cdot u_t, \sigma_f)$, which, given the current state and the velocity, predicts the next state (2D position) of the demonstrator. Predicted positions are therefore corrupted by a Gaussian noise with the fixed variance, σ_f. Without the loss of generality we assume that each inverse model is coupled with the identical forward model. Finally, the observation model is given by a simple model of the motion capture device, $p(z_t|x_t) = \mathcal{N}(x_t, \sigma_o)$, and it provides the prediction error used to drive the recognition process.

Agent's a priori knowledge is represented in the distribution over the contextual variable, $p(c_t)$ which directly biases the choice of internal models through the process of contextual induction $p(i_t|c_t)$ which implicitly encodes the prior knowledge on which actions are most likely to be observed in a given context[2]. For each experiment we compute the posterior distribution $p(i_t|z_{1:t})$ through Monte Carlo integration. The number of particles in all experiments was set to 500.

We recorded 30 video clips at 25 frame-per-second (fps), showing the demonstrator approaching one of several possible objects on a table (average video length was 1,5s). We have divided recordings into four (2x2) groups depending on the number of possible target objects ('simple' group, containing exactly two target objects, and 'complex' group containing up to five target objects) and presence or absence of switching actions indicating novelty ('presence' vs 'absence' groups). Each group contains the same number of recordings. At every frame (40ms) the demonstration was interrupted and we asked participants if they were able to recognize the target of the action corresponding to the goal-directed action (*reaching-object#1*, *reaching-object#2*, ...) by pressing a key on a computer keyboard corresponding to the recognized goal-directed action or an uncertain '*I-dont-know*' response. In order to ease the recognition task we numbered each object in the video. Individual participants ($n = 5$) were randomly selected members of the student population unaware of the purpose of the experiment.

The goal of the system, which plays the role of an "observer", is to infer which of its internal models (i_t) provides the best explanation of the perceived demonstration. The 'response' of the system is measured as its current belief (posterior probability of the winning action), and the model having the probability above a fixed threshold (in the current experiment we empirically set it to 0.7) for at least 200ms (5 frames) is elected as the winning model. The contextual distribution $p(c_t)$ is uniform in all experiments indicating that all objects are equally probable action targets.

We were interested in comparing the instant in which our computational system makes the correct prediction to the instant in which the majority of users recognizes the same goal-directed action. At each frame we measured the average number of correct and wrong/uncertain responses provided by participants. Figure 2 provides a plot of the results in the 'complex/presence' condition; the blue curve depicts the average uncertain (or wrong) user response rate, the green one depicts the positive (e.g. correct) user response rate, while the red curve is the posterior probability of the winning action as computed by our system. Results show that response time our system is qualitatively comparable to that of human participants. In addition, by averaging across several experimental conditions, we have observed that after approximately twenty time steps the number of active models (i.e. those that have associated at least one particle) is only 5% of the total number of models. Note that the recognition of the goal-directed action (measured as the likelihood of the best model) is usually achieved long before the action is terminated, consistent with previously mentioned empirical evidence indicating that action recognition is highly predictive.

To assess the system performance compared to humans', we measured the differences in response times between our system and human participants. For each recording we computed the difference between the instant in which the majority of users recognized the correct action, and the instant in which the posterior probability of the correct internal model was above the recognition threshold for five consecutive frames (200ms). Figure 3 shows the average response time difference with different recognition thresholds in four

[1]The calculated distributions $\mathcal{N}_x$ and $\mathcal{N}_z$ are both assumed to be Gaussian distributions. In our experimental setting this assumption proved to be sufficient having considered unimodal patterns only. In general cases, however, this can lead to poor results and the Gaussian assumption can be relaxed by defining the KL divergence directly over availble particle sets.

[2]This distribution could be easily learned through a life-long supervised inductive mechanism.

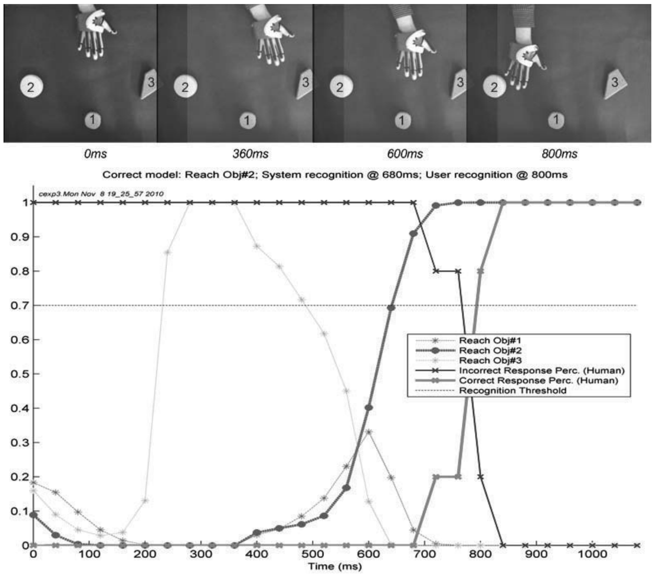


Figure 2: Experimental results: comparing human action recognition with our system in the novelty condition; see text for details.

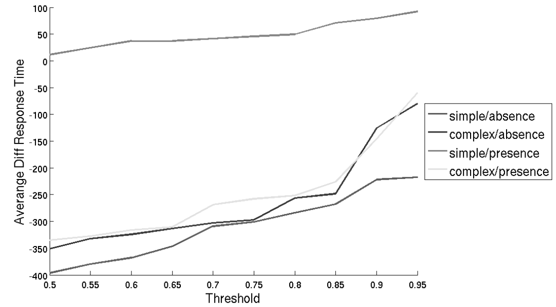


Figure 3: Comparison between user and system response time: effects of the recognition threshold on the response time

experimental settings (corresponding to possible combinations between simple/complex and novelty presence/absence conditions). As shown in the figure, our system responds (on average) faster than human participants, indicating that the proposed method is well suited for real-world scenarios. Interestingly, human users outperform the system in the simple tasks with novel actions. This is due to the fact that, given only two possible target objects, as soon as the novelty is detected humans are certain about the goal of the observed motor act (rather, our system encodes possible goal-directed action alternatives probabilistically).

4 Conclusions and future work

We have proposed that action understanding can be cast as an approximate Bayesian inference, in which inverse-forward models are hypotheses that are tested in parallel via a sampling method, using bounded computational resources. The result is a *motor simulation* process, which is initially biased by prior knowledge of plausible goal-directed actions, and successively refined when novel information is collected that is in agreement or disagreement with the ongoing predictions. In this framework, action goals are implicitly encoded in the coupled inverse-forward models.

In computational motor control, the idea that the brain adopts multiple forward-inverse models to control and recognize actions is quite popular. However, the huge computational complexity of this method prevents its use in most real-world scenarios. Our model is able to efficiently handle the problem of action recognition via a simulative approach based on many internal models and using limited computational resources. This is achieved by adopting an approximate inference procedure (sequential Monte Carlo simulation) for tracking several competing hypotheses. The usage of particle filters in this context is the first attempt - to the best of our knowledge - to provide a computationally efficient implementation of the motor simulation idea, which scales smoothly with the number of available models being the algorithm linear in the (fixed) number of particles. Having more models would require more particles to assure the convergence to the real posterior, but still the algorithm could operate with scarce resources at the cost of reduced accuracy and increasing recognition times.

In addition to its efficiency from an engineering viewpoint, the choice of approximate Bayesian inference could have implications for neuroscience and the problem of what neural coding could support the form of inference that we have argued for. It has been proposed that the brain encodes perceptual and action alternatives probabilistically into populations of neurons [Doya *et al.*, 2007]. Our representational scheme is equivalent to a sampling-based neural representation [Fiser *et al.*, 2010], in which neurons encode variables and their activity at any time representing samples from the distribution of that variable. The hypotheses of reusing one's own motor repertoire during action perception has been tested in several empirical studies which revealed motor involvement during perception [Kilner *et al.*, 2003]. In the cognitive (neuro)science literature, an alternative to the simulation theory is 'theory theory' according to which, to understand another's action, humans rely on an explicit 'theory of mind', or a set of logical hypotheses (possibly expressed in a propositional format) of what causes actions combined with a rationality principle (i.e. what would be rational in a given situation) [Csibra and Gergely, 2007]. In a series of computational studies, [Baker *et al.*, 2006; 2009] describe action understanding as a rational Bayesian process of *inverse planning*, which is close to the idea of "theory theory". A drawback of this method is its computational cost and its huge demands in terms of prior knowledge. Authors use Value Iteration algorithm for estimating a policy in a MDP-equivalent setting which is, at best, quadratic in the number of possible models. In a related approach, [Ramirez and Geffner., 2010] use classical planners rather than Bayesian inference.

Future work will include the extension of our computational model for the recognition of distal intentions through a hierarchy of inverse and forward models in which higher-level pairs encode increasingly abstract actions, in line with the view that mirror neuron mechanism supports the recognition of sequences of actions through dedicated “action chains” and not only individual acts [Fogassi *et al.*, 2005]. Within this architecture, knowledge about plausible distal intentions can serve as priors to recognize proximal actions; conversely, the recognition of proximal actions can serve as prior for inferring the demonstrator’s distal intention.

Acknowledgments

The present study was supported by the EU project “*HUMANOBS: Humanoids That Learn Socio-Communicative Skills Through Observation*” under grant agreement FP7-231453. The effort of Giovanni Pezzulo was also supported by the EU project “*Goal-Leaders*” under grant agreement FP7-270108.

References

[Baker *et al.*, 2006] Chris Baker, Josh Tenenbaum, and Rebecca Saxe. Bayesian models of human action understanding. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Advances in Neural Information Processing Systems 18*, pages 99–106. MIT Press, Cambridge, MA, 2006.

[Baker *et al.*, 2009] Chris L. Baker, Rebecca Saxe, and Joshua B. Tenenbaum. Action understanding as inverse planning. *Cognition*, 113(3):329–349, September 2009.

[Bishop, 2006] C.M. Bishop. *Pattern recognition and machine learning*. Springer New York, 2006.

[Csibra and Gergely, 2007] G. Csibra and G. Gergely. 'obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, 124:60–78, 2007.

[Cuijpers *et al.*, 2006] Raymond H. Cuijpers, Hein T. van Schie, Mathieu Koppen, Wolfram Erlhagen, and Harold Bekkering. Goals and means in action observation: a computational approach. *Neural Networks*, 19(3):311–322, 2006.

[Dearden and Demiris, 2005] A. Dearden and Y. Demiris. Learning forward models for robots. *International Joint Conference on Artificial Intelligence (IJCAI)*, 19:1440, 2005.

[Demiris and Khadhouri, 2006] Y. Demiris and B. Khadhouri. Hierarchical attentive multiple models for execution and recognition of actions. *Robotics and Autonomous Systems*, 54(5):361–369, 2006.

[Doucet *et al.*, 2000] A. Doucet, S. Godsill, and C. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and computing*, 10(3):197–208, 2000.

[Doya *et al.*, 2007] K. Doya, S. Ishii, A. Pouget, and R. P. N. Rao, editors. *Bayesian Brain: Probabilistic Approaches to Neural Coding*. The MIT Press, 1 edition, January 2007.

[Fiser *et al.*, 2010] J. Fiser, P. Berkes, G. Orbán, and M. Lengyel. Statistically optimal perception and learning: from behavior to neural representations. *Trends in Cognitive Sciences*, 14(3):119–130, 2010.

[Fogassi *et al.*, 2005] L. Fogassi, P.F. Ferrari, F. Chersi, B. Gesierich, S. Rozzi, and G. Rizzolatti. Parietal lobe: from action organization to intention understanding. *Science*, 308:662–667, 2005.

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

[Haruno *et al.*, 2001] M. Haruno, D. M. Wolpert, and M. Kawato. Mosaic model for sensorimotor learning and control. *Neural Comput*, 13(10):2201–2220, Oct 2001.

[Kilner *et al.*, 2003] J.M. Kilner, Y. Paulignan, and S.J. Blakemore. An interference effect of observed biological movement on action. *Current Biology*, 13:522–525, 2003.

[Kilner *et al.*, 2007] James M. Kilner, Karl J. Friston, and Chris D. Frith. Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, 8(3), 2007.

[Lopes and Santos-Victor, 2005] M. Lopes and J. Santos-Victor. Visual learning by imitation with motor representations. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 35(3):438–449, 2005.

[Miall and Wolpert, 1996] R. C. Miall and D. M. Wolpert. Forward models for physiological motor control. *Neural Networks*, 9(8):1265–1279, 1996.

[Murphy, 2002] Kevin P. Murphy. *Dynamic bayesian networks: representation, inference and learning*. PhD thesis, UC Berkeley, Computer Science Division, 2002.

[Oztop *et al.*, 2006] E. Oztop, M. Kawato, and M. Arbib. Mirror neurons and imitation: A computationally guided review. *Neural Networks*, 19(3):254–271, 2006.

[Perkell *et al.*, 1997] J. Perkell, M. Matthies, H. Lane, F. Guenther, R. Wilhelms-Tricarico, J. Wozniak, and P. Guiod. Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication*, 22(2-3):227–250, 1997.

[Ramirez and Geffner., 2010] M. Ramirez and H. Geffner. Probabilistic plan recognition using off-the-shelf classical planners. In *Proc. of the AAAI*, Atlanta, USA, 2010.

[Rizzolatti and Craighero, 2004] G. Rizzolatti and L. Craighero. The mirror-neuron system. *Annual Review of Neuroscience*, 27:169–192, 2004.

[Schultz *et al.*, 1997] W. Schultz, P. Dayan, and P.R. Montague. A neural substrate of prediction and reward. *Science*, 275:1593–1599, 1997.

[Wolpert and Kawato, 1998] D. M. Wolpert and M. Kawato. Multiple paired forward and inverse models for motor control. *Neural Networks*, 11(7-8):1317–1329, 1998.

[Wolpert *et al.*, 1995] D. M. Wolpert, Z. Gharamani, and M.I. Jordan. An internal model for sensorimotor integration. *Science*, 269:1179–1182, 1995.

[Wolpert, 2003] D.M. Wolpert. A unifying computational framework for motor control and social interaction. *Philosophical Transactions: Biological Sciences*, 358(1431):593–602, 2003.