If it looks like online control, it is probably model-based control

Dominik Straub (dominikstrb@mailbox.org)
Centre for Cognitive Science, Technical University of Darmstadt
Institute of Psychology, Technical University of Darmstadt

Constantin A. Rothkopf (constantin.rothkopf@tu-darmstadt.de)
Centre for Cognitive Science, Technical University of Darmstadt
Institute of Psychology, Technical University of Darmstadt

Abstract

The interception of moving targets is a fundamental sensorimotor task involving perception and action. For this task, the dominant approach has been to model the behavioral dynamics using online control laws such as the constant bearing angle strategy, which explain behavior without assuming internal models. Here, we derive a Bayesian model-based optimal control model of an interception task and compare it against the constant bearing angle strategy. First, we show that both models equivalently capture average trajectories, suggesting that observing the interception trajectories in an experiment cannot adjudicate between the two models. However, including realistic levels of perceptual uncertainty, motor variability, and sensorimotor delays leads online control without an internal model to quickly deteriorate in interception performance. We conclude that the empirically observed robustness of the constant bearing angle strategy speaks against a direct coupling of environmental variables and behavior, but instead implies some form of internal model.

Keywords: perception and action; online control; model-based control; Bayesian models; dynamical systems

Introduction

Do people employ an internal model of their environment to guide their actions? This is a central question in several areas of cognitive science, particularly in perception and action. The question bears on important theoretical debates in the field, such as the one between perception as inference (Von Helmholtz, 1867; Knill & Richards, 1996) and ecological or direct perception (Gibson, 1979; Warren, 2005), and by extension has been related to the different paradigms of modeling in cognitive science. The Cambridge Handbook of Computational Psychology features Bayesian models of cognition (Griffiths et al., 2008) and dynamical systems approaches to cognition (Schöner, 2008), among others, as two of the most influential paradigms. Discussions regarding these two modeling paradigms have been loaded with broad implications for debates in cognitive science and philosophy of mind about the nature of cognition. Is cognition computational and representational? Or is it embodied, embedded, enacted, and extended?

Besides these perceived implications, both modeling paradigms have been regarded as serving different epistemic goals. Bayesian models have been chosen typically in the rational analysis of a task (Anderson, 1991), describing the goals of a cognitive system. By contrast, dynamical systems approaches have been regarded as the framework of choice for describing the laws underlying behavior and therefore sometimes as an alternative to the dominant paradigm of computational cognitive science (Van Gelder, 1995, 1998). In this view, one could also say that they answer different parts of the questions posed by Marr's (1982) computational level of analysis: dynamical system models are after the descriptive 'what' question, while Bayesian models tend to address the normative 'why' question. Thus, the perceived dichotomy between these modeling frameworks becomes particularly pronounced when interpreting these models as process models or when looking for putative model components at the implementational level. However, although the philosophical positions associated with these two modeling frameworks have often been conceptualized as opposed, the underlying mathematical frameworks themselves are not in conflict (Beer, 1995) and there is diversity in the kinds of explanations obtained from dynamical systems models (Zednik, 2011).

A prime example of a behavior for which there is an ongoing debate between Bayesian and dynamical systems approaches is locomotor interception, i.e. the task of intercepting moving targets. In this field, the contrasting approaches have been called "model-based control" versus "online control" (Zhao & Warren, 2015). In model-based control, "perceptual information and prior knowledge are used to construct and update internal representations or models of the environment, which in turn serve as the basis for control" (Fajen, 2021). The online control approach, on the other hand, posits that "the actor's behavior is primarily determined by currently available perceptual information" (Fajen, 2021). We formalize these notions in a common mathematical framework in the section **Computational modeling**.

Online control strategies for interception tasks have been conceptualized within the framework of dynamical systems under the term *behavioral dynamics* (Fajen & Warren, 2007; Warren & Fajen, 2008). Most famously, using the constant bearing angle (CBA) strategy, an actor maintains a constant angle between the vector from their own position to the target and an external reference direction. Instead of combining the sensory information with internal model predictions, the agent's actions are "coupled" with current sensory input. Specifically, for the interception of moving targets, the relevant information are angular quantities that are thought to be contained in the information from the optic array (Fajen, 2021). In a broader context, Gigerenzer & Brighton (2009)

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

have argued that such simple control strategies, or heuristics, can be more robust and efficient than more complex strategies.

Dynamical systems models are used to model interception tasks because they are well suited to capture the temporal evolution of behavior. Bayesian models, on the other hand, are often thought of as static and suitable for modeling single decisions. However, this gap has been bridged by normative theoretical frameworks for sequential decision-making, such as partially observable Markov decision processes (POMDPs; Åström, 1965; Kaelbling et al., 1998), which can be seen as an extension of Bayesian decision theory to the sequential setting, and the field of optimal control under uncertainty has undergone continuous progress over the last decades. While POMDPs are notoriously hard to solve in general (Papadimitriou & Tsitsiklis, 1987), certain special cases admit tractable approximate solutions. In sensorimotor control, these theoretical ideas have led to the development of stochastic optimal feedback control models (Wolpert & Ghahramani, 2000; Todorov & Jordan, 2002; Shadmehr & Mussa-Ivaldi, 2012).

Although the discussion of computational models so far has excluded specific biological constraints, much is known about the human sensorimotor system. It is well known that human perception is noisy or uncertain, actions underlie structured variability, and that there are temporal delays in the sensorimotor system, which have all been quantified experimentally (Harris & Wolpert, 1998; Yuille & Kersten, 2006; Faisal et al., 2008; Franklin & Wolpert, 2011). Furthermore, achieving task goals involves biomechanical and cognitive effort, which have been formalized as cost functions and also been measured empirically (Körding & Wolpert, 2004; Acerbi et al., 2014). Stochastic optimal control provides a natural way to incorporate these algorithmic level properties into computational models. This makes optimal control under uncertainty a natural framework for building normative models of sequential perception-action tasks, which has found widespread application not only for short reaching and pointing movements (Todorov, 2004) but also for more complex goal-directed behaviors such as ball catching (Belousov et al., 2016) or navigation (Kessler et al., 2022).

Here, we computationally investigate a locomotor interception task as an exemplary behavior for which online control based on dynamical systems has been the dominant modeling framework. To enable a comparison between existing online control strategies like CBA and model-based strategies, we formalize interception tasks in a way that allows for the application of both approaches. Specifically, we derive a non-linear stochastic dynamical system as a description of the agent-environment loop, and highlight how both approaches can be seen in a common mathematical framework. For the model-based agent, we use optimal control methods for non-linear partially observed stochastic systems. Specifically, we approximately solve the estimation and control problem using the extended Kalman filter (EKF) and iterative linear-quadratic Gaussian control (iLQG; Todorov & Li, 2005).

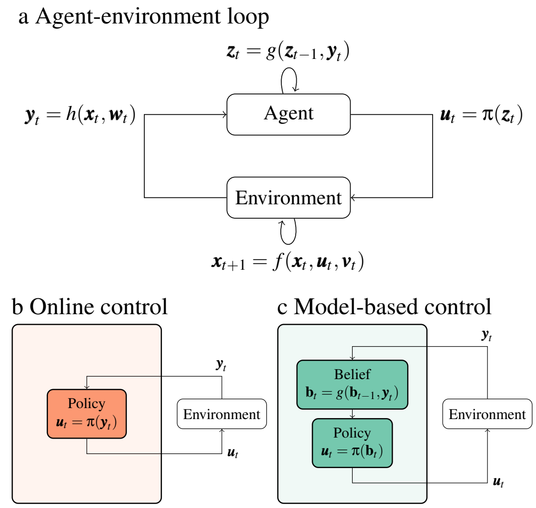

Figure 1: Agent-environment loop. a) The agent receives an observation $\mathbf{y}_t$, which may be a noisy or partial version of the environment state $\mathbf{x}_t$. Based on this observation and potentially an internal state, they perform an action $\boldsymbol{u}_t$, which affects the state at the next time step. b) In online control, the action is directly a function of the observation. c) In model-based control, the agent acts based on an internal state in the form of a belief $\mathbf{b}_t$.

We first show that behavior generated by either model can be closely replicated by fitting it with the respective other model. This works because, for low sensorimotor noise and delay, the behavior of both models can be very similar. Thus, the observation of behavioral dynamics is not sufficient to adjudicate between the two models. However, if biologically realistic perceptual uncertainty, motor variability, and sensory delays are included in the two models, online control quickly breaks down while model-based control stays robust. This suggests that people employ some form of an internal model in order to stably and reliably generate behavioral dynamics.

Computational modeling

In this section, we introduce the necessary mathematical foundation for modeling an agent-environment loop (Fig. 1a), which allows us to implement and compare online and model-based agents. We consider an agent acting in an environment characterized by a discrete-time stochastic non-linear dynamical system:[1]

$$\boldsymbol{x}_{t+1} = f(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{v}_t). \tag{1}$$

The state $\boldsymbol{x}_t$ contains features of the environment such as target positions, but also externally accessible features of the agent such as their position or heading direction. It evolves as a function of the previous state, the action $\boldsymbol{u}_t$ performed by

[1]Continuous-time formulations for online and model-based control can be obtained by replacing the discrete-time difference equations with differential equations.

the agent, and a noise term $\boldsymbol{v}_t \sim \mathcal{N}(0,I)$.[2] This noise term incorporates stochasticity in the dynamics, e.g. in the motor system, but it can also include uncertainty about the internal model. Note that this formulation allows non-normally distributed uncertainty through the function f.

We also want to account for the possibility that not all features of the environment's state are accessible to the agent, but that they instead observe only certain visual variables. We therefore assume that the agent receives an observation of the state at each time step

$$\boldsymbol{y}_t = h(\boldsymbol{x}_t, \boldsymbol{w}_t), \tag{2}$$

which contains any visual variables available to the agent (e.g. distance to the target, bearing angle). It can also incorporate a noise term $\boldsymbol{w}_t \sim \mathcal{N}(0,I)$, modeling uncertainty in the sensory system. Again, as the noise can be transformed through the function h, this formulation allows for non-normally distributed uncertainty, which e.g. can accommodate Weber-Fechner phenomena.

We assume that the agent performs an action according to a policy

$$\boldsymbol{u}_t = \pi(\boldsymbol{z}_t), \tag{3}$$

which is a function of the (internal) state $\boldsymbol{z}_t$. The agent's state evolves as a function of itself and the incoming sensory information,

$$\boldsymbol{z}_t = g(\boldsymbol{z}_{t-1}, \boldsymbol{y}_t). \tag{4}$$

This basic formulation of the agent-environment loop is consistent both with dynamical systems approaches (Beer, 1995; Warren, 2006) and with model-based optimal control approaches (Todorov & Jordan, 2002; Diedrichsen et al., 2010). In the following sections, we flesh out more concretely which assumptions about the specific parts of the agent are made in online and model-based control.

Online control

Online control (Fig. 1b) assumes that the agent acts based solely on the currently available visual information (Zhao & Warren, 2015; Fajen, 2021):

$$\boldsymbol{u}_t = \pi_{\text{online}}(\boldsymbol{y}_t). \tag{5}$$

Importantly, the agent does not need to have a model of the world or form representations of the state of the world. In that sense, the agent's state is given directly by the sensory input, and we arrive at a special case of the general agent-environment loop with $\boldsymbol{z}_t = \boldsymbol{y}_t$. This is sometimes referred to as perceptual variables and actions being "coupled". Note that dynamical systems approaches would in general allow for more complex internal state of the agent, which evolve temporally according to some function, as in Eq. (4). But here, we focus on the online control case, as it has been brought forward in the literature on interception tasks.

[2]The assumption of noise, variability, or uncertainty is much more common in the Bayesian framework than in dynamical systems approaches, but we include it for the sake of generality, because it is well known that the sensorimotor system is subject to variability.

Model-based control

In model-based control (Fig. 1c), the agent combines sensory information with some form of an internal model to mitigate uncertainties and variability. Because the agent does not have access to the true physical state $\boldsymbol{x}_t$, they form a belief about the state of the world based on the internal model using the current observation and previous belief

$$\mathbf{b}_t = g(\mathbf{b}_{t-1}, \boldsymbol{y}_t). \tag{6}$$

The belief is a sufficient statistic of the distribution $p(\boldsymbol{x}_{t+1} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_t)$ and is continuously and dynamically computed using sequential Bayesian updating (see **Extended Kalman filter**). Note that this involves using visual information when available and does not imply a complete general purpose internal model or perfect knowledge about the true state of the environment.

The agent chooses an action based on their belief according to a policy π:

$$\boldsymbol{u}_t = \pi_{\text{optimal}}(\mathbf{b}_t), \tag{7}$$

which is obtained by minimizing a cost function. In order to behave in a Bayes-optimal fashion, the agent takes the actions that minimize the expected cost given their belief about the state of the world

$$\boldsymbol{u}^*_{1:T} = \underset{\boldsymbol{u}_{1:T}}{\arg\min}\, \mathbb{E}_{\boldsymbol{x}_t \mid \mathbf{b}_t}\left[J\left(\boldsymbol{x}_{1:T}, \boldsymbol{u}_{1:T-1}\right)\right]. \tag{8}$$

Note that this allows for modeling internal costs such as biomechanical or cognitive costs thereby accommodating resource-rationality. To summarize, model-based control is a specific special case of the general agent-environment loop, in which the agent's state is a belief state, $\boldsymbol{z}_t = \mathbf{b}_t$, and in which the action $\boldsymbol{u}_t$ is the one that minimizes the expected cost. Thus, model-based control can be seen as the Bayes-optimal strategy in a sequential task. Here, we approximate the solution to the non-linear control problem using the iterative linearization (see **Iterative linear-quadratic Gaussian**).

Extended Kalman filter For linear-Gaussian systems, the Bayes-optimal filter is the Kalman filter (Kalman, 1960). For non-linear dynamics and non-linear observation functions, an approximately optimal Gaussian belief can be computed using the EKF by linearizing the dynamics and observation function around the current belief. It results in the following equation: $\mathbf{b}_{t+1} = f(\mathbf{b}_t, \boldsymbol{u}_t, 0) + K_t(\boldsymbol{y}_t - g(\mathbf{b}_t, 0))$, where K_t is called the Kalman gain and is computed by applying the usual Kalman filter equations to the linearized system.

We additionally allow for the possibility that the agent receives a delayed observation. To model a temporal delay of Δ time steps, we assume that the agent receives an observation $\boldsymbol{y}_t = h(\boldsymbol{x}_{t-\Delta}, \boldsymbol{w}_t)$, which only depends on the state at time $t-\Delta$. This is implemented by augmenting the state vector so that it contains the current state $\boldsymbol{x}_t$ and previous states $\boldsymbol{x}_{t-1}, \ldots, \boldsymbol{x}_{t-\Delta}$ (Izawa & Shadmehr, 2008). Using this formulation, the delay can be taken into account by the EKF. This means that the model-based agent has a model of the system's sensorimotor delay and uses this model to compute beliefs.

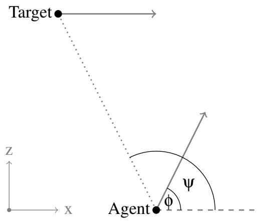

Figure 2: Birds-eye sketch of an interception task. An agent tries to intercept a moving target, while moving in heading direction ϕ. The bearing angle ψ is the angle between the vector from the agent to the target and a reference direction.

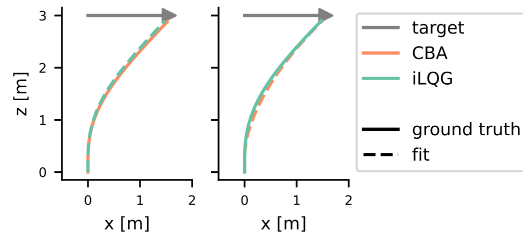

Figure 3: Average trajectory from both agents (online control CBA and model-based iLQG) intercepting a target starting at an initial distance of 3m. In each case, the trajectory from one agent was fit with the other one.

Iterative linear-quadratic Gaussian The control problem defined above can be solved approximately using iLQG (Todorov & Li, 2005), which works by linearizing the dynamics around a nominal trajectory and applying linear-quadratic control (Kalman, 1964) to obtain an optimal policy for the linearized system. This policy is then applied to the full system to yield a new nominal trajectory. Starting with an initial policy, this procedure is iteratively applied until convergence. The result is an approximately optimal linear control law $\pi_t(\mathbf{b}_t) = L_t(\mathbf{b}_t - \bar{\mathbf{x}}_t) + l_t + \bar{\boldsymbol{u}}_{1:T}$, where $(\bar{\boldsymbol{x}}_{1:T}, \bar{\boldsymbol{u}}_{1:T})$ is the nominal trajectory and (L_t, l_t) comprise the control law.

Modeling interception tasks

We model an interception task like the one sketched in Fig. 2. An agent is situated at $\mathbf{p} = \begin{bmatrix} p_x & p_z \end{bmatrix}^T$ and is trying to intercept a target, which is currently at $\mathbf{r} = \begin{bmatrix} r_x & r_z \end{bmatrix}^T$ and moving with velocity $\dot{\mathbf{r}}$. The agent moves with constant speed in a direction given by the heading angle ϕ, which changes with angular velocity $\dot{\phi}$.

The agent controls the heading direction by applying an acceleration $\dot{\phi}_{t+1} = \dot{\phi}_t + dt\mathbf{u}_t + \sigma_m \mathbf{u}_t \mathbf{v}_t$, which is subject to signal-dependent variability. This control can be given by an online control or model-based control policy.

At each time step, the agent receives a partial observation of the state

$$\boldsymbol{y}_t = \bar{\mathbf{y}}_t + \sigma_o \bar{\mathbf{y}}_t \boldsymbol{w}_t, \tag{9}$$

with $\bar{\mathbf{y}}_t = \begin{bmatrix} \dot{\Psi} & D & \dot{\Phi} \end{bmatrix}^T$. The observation contains the bearing angle's velocity $\dot{\Psi}$, the Euclidean distance between agent and target $D = \|\mathbf{p} - \mathbf{r}\|$, and the heading velocity $\dot{\Phi}$. This choice is motivated by the variables that are necessary for the CBA strategy and which have been shown empirically to be accessible to humans (Lenoir et al., 1999). We denote the noisy versions of these variables using lower-case letters $\mathbf{y}_t = \begin{bmatrix} \dot{\psi} & d & \dot{\phi} \end{bmatrix}^T$. Because the noise term $\boldsymbol{w}_t$ is multiplied by $\bar{\boldsymbol{y}}_t$, the observation is subject to signal-dependent noise, implementing Weber's law in the observation function.

Model-based control cost function For the model-based agent, we model the agent's goal in the interception task using the following cost function:

$$J(\boldsymbol{x}_{1:T}, \boldsymbol{u}_{1:T-1}) = \|\mathbf{p}_T - \mathbf{r}_T\|^2 + \alpha \sum_{t=1}^{T-1} \|\boldsymbol{u}_t\|^2. \tag{10}$$

It consists of two terms. First, the agent wants to reach the target at the final time step, i.e. minimize the (quadratic) distance to the target. Second, the agent minimizes the control effort expended over the whole trajectory. This second term is weighted by a free parameter α, which trades off the main goal of reaching the target with the penalty on action effort.

Constant bearing angle strategy The most popular online control strategy that has been proposed in the context of locomotor interception in humans and animals is the constant bearing angle (CBA) strategy (Fajen & Warren, 2007). It can be expressed as a dynamical system describing the relationship between sensory input and the action performed by the agent. The CBA strategy is defined as

$$\mathbf{u}_t = -b\dot{\phi}_t + k\dot{\psi}_t(d_t + c). \tag{11}$$

It consists of a damping term proportional to the current rate of change of the heading direction $\dot{\phi}_t$ and a term responsible for bringing the rate of change of the bearing angle $\dot{\psi}_t$ to zero, which is scaled by the distance to the target d_t. Following Fajen & Warren (2007), we fixed $c = 1.0$, which results in two free parameters: the damping parameter b and the stiffness parameter k.

Results

To compare the online and model-based agents introduced above, we use a locomotor interception task closely based on the one used in the experiment by Fajen & Warren (2004). This task has previously been modeled using online control by fitting the CBA policy's free parameters to human empirical data (Fajen & Warren, 2007). In the experiment, participants walked freely in a 12 by 12 m area. They were presented with a virtual environment using a head-mounted display. In each trial, a target appeared at a distance of 3 m and participants were instructed to walk to the target, which moved at a constant speed of 0.6 m / s. The agent's movement speed was set to 1.29m/s, following Fajen & Warren (2007).

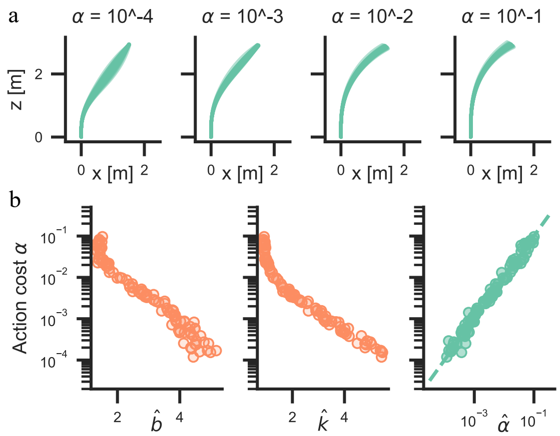

Figure 4: a) Model-based control trajectories for different control cost parameters α. b) For data simulated from iLQG with different values of α, we obtained parameter estimates ($\hat{b}$, $\hat{k}$) by fitting the CBA strategy. We then simulated trajectories using these estimated parameters and fit iLQG to these to obtain estimates of the control cost $\hat{\alpha}$.

We simulated a model-based control agent using the EKF and iLQG and an online control agent using the CBA strategy. For these initial simulated experiments, we set the sensorimotor delay to 0 ms and sensorimotor noises to relatively low levels ($\sigma_m = 0.05, \sigma_o = 0.25$). In addition to these sensorimotor characteristics, both models have free parameters that determine their action policies. Model-based iLQG control is based on optimizing a cost function. In our interception model, the cost function has one free parameter α, which affects how strongly controls are penalized. CBA has two free parameters: the damping parameter b and the stiffness parameter k. We set the parameters k and b of CBA to those values that were reported to match human behavior (from Fajen & Warren, 2007) and the action cost parameter for iLQG to $\alpha = 1e-3$. Fig. 3 shows the average trajectory from both agents with these settings. Both the model-based iLQG (left column) and the online CBA turn right early in the trajectory and then walk on a relatively straight path until they intercept the target. We then fit the online control agent to the behavior from the model-based agent and vice versa using non-linear least squares. In both cases, we obtain a good fit that closely reproduces the other agent's behavior when simulated with the best-fitting parameters.

This works not only for one parameter setting, but across a wide range of plausible parameter values. Varying the action cost parameter of the iLQG agent leads to trajectories with different curvature without affecting interception success (Fig. 4a). We fit the online control agent to the simulated model-based behavior for a range of action costs between $\alpha = 10^{-4}$ and $\alpha = 10^{-1}$ sampled uniformly in logarithmic space. For each value of α, we thereby obtained estimates of the online control agent's free parameters. The estimated parameters $\hat{b}$ and $\hat{k}$ of CBA change as a function of the action cost parameter (Fig. 4b): with increasing action cost, we see

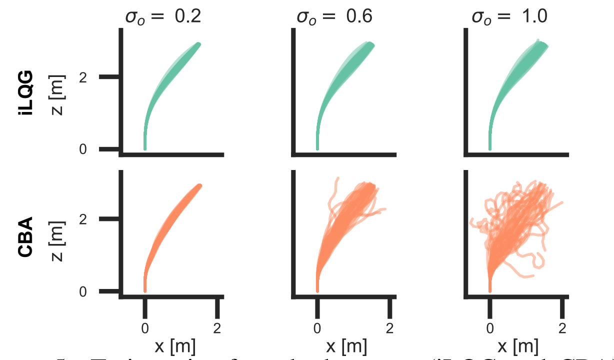

Figure 5: Trajectories from both agents (iLQG and CBA) with a sensorimotor delay of 100 ms intercepting a target starting at an initial distance of 3m.

an increase in both the damping $\hat{b}$ and the stiffness $\hat{k}$.

Using each pair of estimated CBA parameters, we again simulated trajectories of online control agents. To each of these sets of trajectories, we fit the model-based strategy. The resulting estimates $\hat{\alpha}$ closely match the original values of α, from which the CBA parameters were previously obtained (Fig. 4, right subplot). Thus, based on these simulations, we conclude that both the online control and the model-based control formulations are equivalent in capturing average observed interception trajectories. Accordingly, we cannot adjudicate between the two models on the basis of observed average trajectories.

While the above simulations were agnostic to the known physiological uncertainties and delays in the human sensorimotor system, we now turn to incorporating these constraints into the models. Fig. 5 shows simulations from both models with a sensorimotor delay of 100 ms, motor noise of $\sigma_m = 0.1$ and three different levels of observation noise ($\sigma_o \in \{0.2, 0.6, 1.0\}$).[3] At low levels of sensory noise, the trajectories from iLQG (top row) and CBA (bottom row) are very similar and indistinguishable with the bare eye. However, while iLQG remains robust to increasing levels of sensory noise, CBA becomes more variable, with several trials veering off the path towards the target and failing to intercept it successfully.

We investigated this in more detail by simulating different temporal delays (100, 200ms) and a range of different observation noises (σ_o from 0.04 to 1.0) and motor noises (σ_m from 0.02 to 0.5), all within physiologically plausible ranges (Kessler et al., 2022). We defined a successful interception as a trial in which the final distance from the agent to the target was less than 75 cm. This can be interpreted as the agent being able to reach the target with an extended arm. The results are summarized in Fig. 6. The model-based agent (iLQG) is robust against sensorimotor noise and delays (left column), with success rates not falling below 0.75. The online control

[3] We chose the maximum level of $\sigma_o = 1.0$ based on the Weber fractions for visual speed discrimination measured by Hogendoorn et al. (2017), which we adjusted to account for the length of the time steps in our simulations.

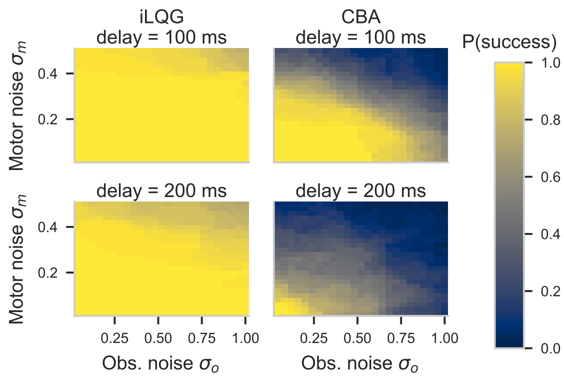

Figure 6: Interception success rate for both strategies and multiple levels of delays, observation noise and motor noise.

agent (CBA) starts breaking down when the scaling of sensory noise or motor noise minimally increases. This effect is even stronger at larger temporal delays.

Discussion

The debate between model-based and model-free accounts of perception and action is important under different names in several fields of cognitive science. One exemplary behavior is locomotion. Extensive experiments have been conducted, in which human and animal behavior has been collected and modeled with the CBA strategy (e.g. Lenoir et al., 2002; Ghose et al., 2006; Fajen & Warren, 2007). However, essentially identical trajectories as those from CBA can also arise from a model-based control agent. As we have demonstrated in simulations, it is possible to fit a dynamical system to the data from a model-based agent with different behavioral costs. Thus, finding a dynamical system that reproduces the observed trajectories is not sufficient to claim that the behavior did in fact arise without an internal model. Of course, this does not prove that people in fact use model-based strategies in interception tasks, but it should caution against inferences about the nature of cognition based on being able to account for observed behavior with a dynamical system.

What the simulations presented here also illustrate is that the two modeling approaches serve different epistemic goals: rational models, like Bayesian optimal control models, provide normative explanations, while dynamical systems aim at a descriptive account. In the interception task, all parameters are cognitively interpretable and can be related to empirical research in psychology or neuroscience, e.g. the motor variability or perceptual uncertainty. The cost parameter of the model-based agent offers an intentional explanation for the agent's behavior. They acted in a certain way because of a subjective trade-off between task goals and intrinsic costs of behavior, which could be physiological (Di Prampero, 1986) or cognitive costs (Shenhav et al., 2017) and which aligns with the notion of resource rational behavior (Simon, 1955; Anderson, 1991; Gershman et al., 2015). This is different from the kind of explanation we obtain from the dynamical system. The damping and stiffness parameters are difficult to relate to an agent's goals or desires, but rather serve to provide a law-like description of behavior.

It is also worth noting that model-based control can be extended in a straightforward way to incorporate additional cognitive constraints. Optimal control under uncertainty can depart from ideal observer models (Geisler, 1989), which assume perfect knowledge of the generative model of the world, by incorporating possibly biased or false beliefs about dynamics. Estimating these beliefs from behavior using inverse optimal control reconciles descriptive and normative models (Rothkopf & Ballard, 2013; Wu et al., 2020; Straub & Rothkopf, 2022). Rather than positing costs and beliefs and comparing the predictions of a normative model against behavior, costs and beliefs are treated as latent variables that are inferred from behavior with a generative model. Thus, these methods answer what costs and beliefs the behavior is implicitly optimal for. Moreover, this framework allows quantifying the uncertainty of the researcher about the inferred cognitive quantities on the basis of empirical data.

While behavioral studies on interception typically ignore behavioral variability and focus on average trajectories per condition (e.g. Fajen & Warren, 2004; Zhao et al., 2023), our simulations show that looking at variability is critical to distinguish different strategies. We investigated sources of behavioral variability that are known to affect human sensorimotor behavior: sensorimotor delay (Crevecoeur & Gevers, 2019) and signal-dependent noise in controls (Harris & Wolpert, 1998; Jones et al., 2002). Different models showed different increases of behavioral variability with increasing noise and delay. Specifically, model-based control was more robust against delay and noise than online control. Thus, considering behavioral variability seems a promising way to adjudicate between competing theories. One way to achieve this is to use computational models and statistical methods that model the characteristics of the sensorimotor system and estimate them from behavioral variability (Schultheis et al., 2021). Future work should test the implications of the modeling results empirically.

From a broader perspective, our simulation results are closely related to the idea that in order to be a good controller of a system, one needs to have a model of the system (Francis & Wonham, 1976). However, those results relate to deterministic systems, while the present simulations and result include sensorimotor noises and delays.

Finally, we do not argue that there is no value in deriving law-like descriptions of human behavior in the form of dynamical systems. The dynamical approach to cognition has the benefit of emphasizing the temporal structure of behavior and the interaction between the body and the environment, an aspect that has often been missing in previous Bayesian models. However, advances in optimal control under uncertainty have enabled the application of normative (Bayesian) ideas to sequential decision-making, allowing for a top-down research strategy that descends Marr's levels (Zednik & Jäkel, 2016).

Acknowledgments

We would like to thank Carlos Zednik, Fabian Tatai, and Inga Ibs for feedback on an earlier draft, and Matthias Schultheis for his help with developing the optimal control code framework. We gratefully acknowledge the computing time provided to us on the high-performance computer Lichtenberg at the NHR Centers NHR4CES at TU Darmstadt. This research was supported by the European Research Council (ERC; Consolidator Award 'ACTOR'-project number ERC-CoG-101045783).

References

Acerbi, L., Ma, W. J., & Vijayakumar, S. (2014). A framework for testing identifiability of bayesian models of perception. *Advances in neural information processing systems*, *27*.

Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral and brain sciences*, *14*(3), 471–485.

Åström, K. J. (1965). Optimal control of Markov processes with incomplete state information I. *Journal of Mathematical Analysis and Applications*, *10*, 174–205.

Beer, R. D. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial intelligence*, *72*(1-2), 173–215.

Belousov, B., Neumann, G., Rothkopf, C. A., & Peters, J. R. (2016). Catching heuristics are optimal control policies. *Advances in neural information processing systems*, *29*.

Crevecoeur, F., & Gevers, M. (2019). Filtering compensation for delays and prediction errors during sensorimotor control. *Neural computation*, *31*(4), 738–764.

Diedrichsen, J., Shadmehr, R., & Ivry, R. B. (2010). The coordination of movement: optimal feedback control and beyond. *Trends in cognitive sciences*, *14*(1), 31–39.

Di Prampero, P. (1986). The energy cost of human locomotion on land and in water. *International journal of sports medicine*, *7*(02), 55–72.

Faisal, A. A., Selen, L. P., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature reviews neuroscience*, *9*(4), 292–303.

Fajen, B. R. (2021). *Visual control of locomotion*. Cambridge University Press.

Fajen, B. R., & Warren, W. H. (2004). Visual guidance of intercepting a moving target on foot. *Perception*, *33*(6), 689–715.

Fajen, B. R., & Warren, W. H. (2007). Behavioral dynamics of intercepting a moving target. *Experimental Brain Research*, *180*(2), 303–319.

Francis, B. A., & Wonham, W. M. (1976). The internal model principle of control theory. *Automatica*, *12*(5), 457–465.

Franklin, D. W., & Wolpert, D. M. (2011). Computational mechanisms of sensorimotor control. *Neuron*, *72*(3), 425–442.

Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological review*, *96*(2), 267.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278.

Ghose, K., Horiuchi, T. K., Krishnaprasad, P., & Moss, C. F. (2006). Echolocating bats use a nearly time-optimal strategy to intercept prey. *PLoS biology*, *4*(5), e108.

Gibson, J. J. (1979). The ecological approach to visual perception.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in cognitive science*, *1*(1), 107–143.

Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. In R. Sun (Ed.), *The cambridge handbook of computational psychology* (p. 59–100). Cambridge University Press. doi: 10.1017/CBO9780511816772.006

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, *394*(6695), 780–784.

Hogendoorn, H., Alais, D., MacDougall, H., & Verstraten, F. A. (2017). Velocity perception in a moving observer. *Vision Research*, *138*, 12–17.

Izawa, J., & Shadmehr, R. (2008). On-line processing of uncertain information in visuomotor control. *Journal of Neuroscience*, *28*(44), 11360–11368.

Jones, K. E., Hamilton, A. F. d. C., & Wolpert, D. M. (2002). Sources of signal-dependent noise during isometric force production. *Journal of neurophysiology*, *88*(3), 1533–1544.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial intelligence*, *101*(1-2), 99–134.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, *82*(1), 35-45.

Kalman, R. E. (1964). When Is a Linear Control System Optimal? *Journal of Basic Engineering*, *86*(1), 51-60.

Kessler, F., Frankenstein, J., & Rothkopf, C. A. (2022). A dynamic bayesian actor model explains endpoint variability in homing tasks. *bioRxiv*, 2022–11.

Knill, D. C., & Richards, W. (1996). *Perception as bayesian inference*. Cambridge University Press.

Körding, K. P., & Wolpert, D. M. (2004). The loss function of sensorimotor learning. *Proceedings of the National Academy of Sciences*, *101*(26), 9839–9842.

Lenoir, M., Musch, E., Janssens, M., Thiery, E., & Uyttenhove, J. (1999). Intercepting moving objects during self-motion. *Journal of Motor Behavior*, *31*(1), 55–67.

Lenoir, M., Musch, E., Thiery, E., & Savelsbergh, G. J. (2002). Rate of change of angular bearing as the relevant property in a horizontal interception task during locomotion. *Journal of motor behavior*, *34*(4), 385–401.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman.

Papadimitriou, C. H., & Tsitsiklis, J. N. (1987). The complexity of markov decision processes. *Mathematics of operations research*, *12*(3), 441–450.

Rothkopf, C. A., & Ballard, D. H. (2013). Modular inverse reinforcement learning for visuomotor behavior. *Biological cybernetics*, *107*, 477–490.

Schultheis, M., Straub, D., & Rothkopf, C. A. (2021). Inverse optimal control adapted to the noise characteristics of the human sensorimotor system. *Advances in Neural Information Processing Systems*, *34*, 9429–9442.

Schöner, G. (2008). Dynamical systems approaches to cognition. In R. Sun (Ed.), *The cambridge handbook of computational psychology* (p. 101–126). Cambridge University Press. doi: 10.1017/CBO9780511816772.007

Shadmehr, R., & Mussa-Ivaldi, S. (2012). *Biological learning and control: How the brain builds representations, predicts events, and makes decisions*. MIT Press.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual review of neuroscience*, *40*, 99–124.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 99–118.

Straub, D., & Rothkopf, C. A. (2022). Putting perception into action with inverse optimal control for continuous psychophysics. *Elife*, *11*, e76635.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature neuroscience*, *7*(9), 907–915.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature neuroscience*, *5*(11), 1226–1235.

Todorov, E., & Li, W. (2005). A generalized iterative lqg method for locally-optimal feedback control of constrained nonlinear stochastic systems. In *Proceedings of the 2005, american control conference, 2005.* (pp. 300–306).

Van Gelder, T. (1995). What might cognition be, if not computation? *The Journal of Philosophy*, *92*(7), 345–381.

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and brain sciences*, *21*(5), 615–628.

Von Helmholtz, H. (1867). *Handbuch der physiologischen optik*. Voss.

Warren, W. H. (2005). Direct perception: The view from here. *Philosophical Topics*, *33*(1), 335–361.

Warren, W. H. (2006). The dynamics of perception and action. *Psychological review*, *113*(2), 358.

Warren, W. H., & Fajen, B. R. (2008). Behavioral dynamics of visually guided locomotion. *Coordination: neural, behavioral and social dynamics*, 45–75.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature neuroscience*, *3*(11), 1212–1217.

Wu, Z., Kwon, M., Daptardar, S., Schrater, P., & Pitkow, X. (2020). Rational thoughts in neural codes. *Proceedings of the National Academy of Sciences*, *117*(47), 29311–29320.

Yuille, A., & Kersten, D. (2006). Vision as bayesian inference: analysis by synthesis? *Trends in cognitive sciences*, *10*(7), 301–308.

Zednik, C. (2011). The nature of dynamical explanation. *Philosophy of Science*, *78*(2), 238–263.

Zednik, C., & Jäkel, F. (2016). Bayesian reverse-engineering considered as a research strategy for cognitive science. *Synthese*, *193*, 3951–3985.

Zhao, H., Straub, D., & Rothkopf, C. A. (2023). People learn a two-stage control for faster locomotor interception. *Psychological Research*, 1–20.

Zhao, H., & Warren, W. H. (2015). On-line and model-based approaches to the visual control of action. *Vision research*, *110*, 190–202.