Towards an Anticipatory Mechanism for Complex Decisions in a Bio-Hybrid Beehive

Heinrich Mellmann and Volha Taliaronak and Verena V. Hafner

Abstract To a certain extent, humans and many other biological agents are able to anticipate the consequences of their actions and adapt their decisions based on available information on current and future states of their environment. The same principle can be applied to enable decision-making in artificial agents. In order to decide on an action, an agent could envision the consequences for each of the actions and then choose the one promising the best outcome. This anticipatory scheme can enable fast decisions in highly dynamic and complex situations, which has been demonstrated in humanoid robots playing soccer. We extend this principle to the scenario of bio-hybrid beehives augmented with robotic actuators, which allow to influence the foraging locations of the bees. We investigate how a bio-hybrid beehive can make decisions and direct the bees in a way which would benefit the the whole ecosystem enabling sustainable beekeeping. We explore the general principles of anticipation and discuss connections to cognitive science and developmental robotics. We present an implementation of a simulator for the behavior of the augmented beehive and present preliminary results demonstrating the feasibility of the anticipatory approach.

Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany, e-mail: {mellmann, taliarov,hafner}@informatik.hu-berlin.de
http://hu.berlin/adapt

This is a preprint of the following chapter: Heinrich Mellmann, Volha Taliaronak, Verena V. Hafner, Towards an Anticipatory Mechanism for Complex Decisions in a Bio-Hybrid Beehive, published in Concurrency, Specification and Programming: Revised Selected Papers from the 29th International Workshop on Concurrency, Specification and Programming (CS&P'21), Berlin, Germany, edited by Schlingloff, Bernd-Holger and Vogel, Thomas and Skowron, Andrzej, 2023, Springer reproduced with permission of Springer Nature Switzerland AG. The final authenticated version is available online at: http://dx.doi.org/10.1007/978-3-031-26651-5_7

1 Introduction

With *Bio-Hybrid Beehive* we refer to a symbiotic system consisting of a beehive augmented with robotic components – sensors, actuators and computational resources. These augmentations provide the bio-hybrid beehive with information regarding the state of the colony and the environment, and enables it to influence foraging locations of its bees. From that perspective we can consider the *Bio-Hybrid Beehive* to be a robot that is able to perceive its environment and send its bees to certain locations.

This ability can be used to direct bees to specific areas for foraging with higher yield, and prevent them from foraging in hazardous regions or wildlife preserves protecting wild bees. This could improve the well-being of the honeybees, increase the value generated by them, and reduce the competition with wild bee species. This, in turn, could improve the symbiotic relationship between the honeybees, humans and the ecosystem as a whole. To realize this, the *Bio-Hybrid Beehive* requires an appropriate decision process to decide when and where to send the bees. The foraging behavior of the bees is highly complex and is influenced by a wide variety of factors. Many environmental factors need to be taken into account, including complex topography and actions of other beehives. An action of directing bees to a particular location involves significant and complex uncertainty. This makes inferring a decision a challenging task.

To a certain extent, humans and many other biological agents are able to anticipate the consequences of their actions and adapt their decisions based on available information on current and future states of their environment. This makes anticipation a powerful mechanism enabling complex decision-making and behavior, which should also hold true for artificial agents that interact in complex environments. In an artificial agent, e.g., a robot, anticipation can be realized with the help of an internal simulation. To decide on an action, the internal simulation can be used to envision the outcome of available actions and select the one with the most promising result.

An intuitive approach to making decisions in complex scenarios based on anticipation and internal simulation was presented in [32] and [33] in the scenario of humanoid robot soccer - RoboCup. There, humanoid robots play soccer autonomously and have to perceive the environment, make decisions and execute actions in real time in a highly dynamic and complex environment of a soccer game. The overall decision process can be split into three basic steps: *predict* the possible outcomes of the available actions; *evaluate* the outcome according to desired criteria; and *select* the action promising the highest value. In the case of robot soccer - we could simulate where the ball would land after different possible kick actions and choose the one where the ball lands as close to the opponent goal as possible. The central part of this scheme is the prediction step. To predict the results of an action we can use a physical simulator calculating the behavior of the ball after the kick and its interactions with the environment. This simulator is then used as a part of the decision process and is referred to as *internal forward simulation*.

The anticipatory decision scheme provides a direct approach to address the complexity of the *Bio-Hybrid* scenario, similar to the soccer scenario described above. The general intuition remains the same: we envision what would happen when the bees are directed to a certain location and choose the location promising the most desirable outcome. For example, to maximize the yield of honey, we predict how much honey will be collected for each possible location that the bees can be directed to, and then we select the location promising the highest yield.

To study decision-making in the context of bio-hybrid beehives we develop a simulation for the foraging behavior of the bees. This simulation can be used as both an environment to study decision mechanisms, and an internal simulator in an anticipatory decision scheme to predict the outcome of possible actions. Preliminary work on such simulation and decision making in *bio-hybrid* beehives was published in [51]. In this work we extend the simulation, formulate experimental scenarios for studying decision mechanisms, and discuss the generalization of the anticipatory approach to those scenarios.

We begin with a literature review of studies involving anticipation in robots in Section 2. In Section 3, we discuss the general principle of anticipation and its role in decision-making in artificial agents. We expand and generalize the framework for a decision mechanism based on anticipation from [32] and [33]. In Section 4, we discuss how the principle of anticipation can be applied to realize decision-making in the scenario of a bio-hybrid system consisting of a beehive augmented with technology, which allows it to influence the foraging locations of its bees and provides access to global knowledge, like safe foraging locations or weather forecasts.

2 Literature Review

Often the terms *prediction* and *anticipation* are being used interchangeably. One of the first widely cited characterizations of an anticipatory system was given by Robert Rosen in (Rosen, 1979) [44, p. 537] and later defined more formally in (Rosen, 1985) [45, p. 313]:

> I have come to believe that an understanding of anticipatory systems is crucial, not only for biology, but also for any sphere in which decision making based on planning is involved. These are systems which contain predictive models of themselves and their environment, and employ these models to control their present activities.
>
> (Rosen, 1979) [44, p. 537]

To predict means to envision the future based on the current knowledge of the situation, including possible actions taken by the actors involved. Anticipation, on the other hand, means choosing an action based on the envisioned future. In a certain sense, the mechanism of anticipation can be seen as an *inverse* of the mechanism of prediction.

The concept of anticipation has been extensively studied in different scientific disciplines. The following two collections provide a comprehensive overview – *The Challenge of Anticipation: A Unifying Framework for the Analysis and Design of Artificial Cognitive Systems* [39] and *Anticipatory Behavior in Adaptive Learning Systems: From Psychological Theories to Artificial Cognitive Systems* [40]. A brief overview of anticipatory mechanisms in humans and animals, as well as artificial agents, can be found in [34].

More generally, anticipation appears in a wide variety of control and decision algorithms. For example, an algorithm playing a game of chess might predict several opponent's steps in the future and execute its action in accordance with those predictions. Although the principle of anticipation is very general, we will focus on artificial agents with a physical body, also called *embodied artificial agents*.

Winfield and Hafner [54] consider anticipation in embodied artificial agents through the mechanism of *predictive internal models*, which generate a prediction of a particular state in the future. Depending on the scenario, these models can generate predictions regarding the agent's own body, behavior of other agents (humans or robots), and the environment. Such models can be learned (acquired) by the agent or predefined. Both the fixed and the adaptive model can be used to generate a prediction of a particular state in the future.

A *predictive internal model* is a model of the agent's body, its environment and other agents, which is internal to the agent (part of its cognitive process), and which can be used to make decisions on predictions of the future. In essence, this model needs to capture the laws of reality to a sufficient degree. A predictive internal model can be implicit or explicit. An implicit predictive model can allow reasoning about the future without explicitly computing the state of the future. As an example consider a probabilistic predictive model, which computes the likelihood of occurrence for a given future state. An *internal simulation* is a predictive internal model that generates explicit predictions of the future by simulating real phenomena.

In the following we give a brief overview over studies of anticipatory mechanisms in artificial agents. The studies are roughly divided in two sections: the studies in the Section 2.1 focus on how the anticipatory mechanism can be acquired (learned) by the agent similar to humans and other animals; the studies in the Section 2.2 use anticipation as a tool, and investigate how it can be used as a mechanism to realize complex decisions in realistic scenarios.

2.1 Adaptive Internal Models

Humans and animals are able to acquire predictive models from experience through exploration in the infant developmental phase, and maintain them over time. In the field of developmental robotics, researchers borrow from theories of human and animal development. They study how a robot can acquire and maintain a model of its own body and the environment by learning from experience that is collected through exploration in a similar way to infants. The learning can involve predictive models as well as corresponding inference mechanisms to derive an action based on predictions. The predictive models can be learned completely or in parts, and can represent the own body, the consequences of own actions, the dynamics of the environment, or actions of the others.

In robotics, the ability to acquire a predictive model of the own body and to anticipate has been demonstrated in several studies and experiments. One of the earliest implementations in real hardware was presented in [7] by Bongard, Zykov and Lipson. There, a four-legged robot resembling a starfish was able to simulate its locomotion and was therefore able to adapt to changes in the morphology of its body, such as removed leg parts.

Demiris and Khadhouri proposed an architecture based on hierarchical networks of inverse and forward models introduced in [11]. This architecture is able of selecting and executing an action; as well as perceiving it when performed by a demonstrator. The idea of inverse and forward models was adopted by Schillaci and colleagues in [47], where they demonstrated that a humanoid robot is able to learn internal models of sensorimotor relationships through an exploratory phase inspired by infants' body babbling. The acquired models can be used for decision-making in tool-use scenarios, as has been demonstrated in [46].

Matsumoto and Tani [31] use predictive coding and active inference for goal-directed planning of grasping trajectories with only partial knowledge. In [20, 35, 38] information theoretic measures are used to learn a self-model of a robot from experience.

Anticipation has also been studied in scenarios involving human-robot interaction, where the robot was required to anticipate human behavior. In [15], the authors implement the ability to read human intentions in a humanoid robot *iCub*. In [12], *iCub* acquires multi-modal models of collaborative action primitives, which enable it to recognize intended action of a human based on human's gaze.

Pico and collages have demonstrated in [41] that forward models can be used to predict the noise a wheeled robot produces by intended motor actions. The models are learned in a non-supervised manner based on random exploration of the noise produced by the motors, also called random motor babbling. A comparison between the predicted and perceived noise allows the robot to infer information about its environment.

In a purely simulated study [28], humanoid agents learn to play soccer, whereby all aspects are learned in a multi-stage process, beginning with skills to strategic decisions. The results of the analysis show, in particular, that the agents acquire an ability to predict the behavior of opponents and teammates. The analysis also shows a positive correlation between higher performance of an agent and its ability to predict future game states.

In general, robots' ability to acquire a predictive model by itself is an extraordinarily complex task. In most scenarios, the learning requires a large amount of data and an extensive exploration phase. To manage the complexity, in most scenarios discussed above, learning is limited to parts of the model relying on classical approaches, for instance, for object detection in images.

2.2 Inference of Decisions Based on Anticipation

To realize a decision mechanism based on anticipation, we essentially need a predictive model and an inference mechanism, which uses the predictive model to make decisions. In practical scenarios, we can use external tools to construct a predictive model. For instance, some approaches use a physical simulator calculating the state of the situation based on physical laws. Such internal forward simulation has already been successfully used as an inference method in robotics.

In a minimalist experiment in [6], a small wheeled robot traverses a corridor with moving obstacles. An agent equipped with anticipatory behavior has shown higher success in avoiding collisions in comparison to reactive agents. In [8], the authors investigate the navigation of wheeled robots in a dynamic environment. They use a simulation approach to envision movements of other agents and pedestrians to avoid dynamic obstacles while moving towards a goal.

In [16] a robot equipped with a soft hand explores objects by moving them. The authors use physical simulation of the interaction between the soft hand of the robot and the manipulated object to predict resulting movement of the object. Predictions are used to select actions maximizing the information about the object's properties. In [25] the authors introduce a pancake-baking robot that plans its actions using a full physical simulation of the outcome of possible actions.

There have been several approaches within the RoboCup community to implement decision mechanisms based on anticipation. Mellmann, Schlotter and Blum formulate in [32, 33] an intuitive scheme based on forward simulation, which allows for fast and robust selection of kick-actions in soccer-playing robots. A number of other works focus on a similar task of selecting an optimal kick-action. In [13], a probabilistic approach is used to describe the kick selection problem which is then solved using the Monte Carlo simulation. In [19], the kick is chosen to maximize a proposed heuristic *game situation score*, which reflects the goodness of the situation. In [1], the authors use an instance based representation for the kick actions, and employ Markov decision process as an inference method. In [36], the authors find that projection of the intention of other players can significantly improve the performance of path-planning algorithms.

3 Anticipation and Decision-Making in Artificial Agents

In this section, we discuss how anticipation can be realized in embodied artificial agents. For this, we expand and generalize the framework for a decision mechanism based on anticipation from [32] and [33] by Mellmann and Schlotter, which was used to select kick actions in the scenario of humanoid robot soccer - RoboCup. There, humanoid robots play soccer autonomously and have to perceive the environment, make decisions and execute actions in real time in a highly dynamic and complex environment of a soccer game.

Selecting a kick action can present a significant challenge in a complex situation of a soccer game. The outcome of a particular action may depend on a wide variety of environmental factors, such as the robot's position on the field or the location of other players. In addition, the robot's perception of the situation is often uncertain, noisy and incomplete, and the execution of the actions is subject to noise and uncertainty as well. The anticipatory decision-making scheme presented in [32, 33] provides an intuitive and versatile approach to deal with this complexity.

In general, an anticipatory system is a system that makes decisions based on a prediction of the future. This implies the need for two components: a mechanism for predicting the future, and one to infer decisions based on those predictions. From this we can devise a decision scheme consisting of three basic steps: *predict* the possible outcomes of the available actions; *evaluate* the outcome according to desired criteria; and *select* the action promising the highest value. In the case of robot soccer - we could simulate where the ball would land after different possible kick actions and choose the one where the ball lands as close to the opponent goal as possible. The central part of this scheme is the prediction step. To predict the results of an action we can use a physical simulator calculating the behavior of the ball after the kick and its interactions with the environment. This simulator is then used as a part of the decision process and is referred to as *internal forward simulation*.

In the following two sections, Section 3.1 and Section 3.2, we discuss a modular architecture of the overall cognition and perception process, and practical considerations arising from an anticipatory decision mechanism. In the second half, in the Section 3.1, we discuss a possible formalization for a general anticipatory decision scheme.

3.1 Perception and Decision-Making in Robotics

In this section, we discuss how the cognitive processes and the decision-making processes are commonly realized in complex real world scenarios.

It is challenging to formulate the complete cognitive process of a robot in a real world scenario as a single monolithic process, for instance, as a stochastic process or a neural network. An example of this is the Partially observable Markov decision process (POMDP). POMDP is a general and powerful framework used to formulate the decision process for an agent in a partially observable environment. It works well in robots in isolated real world scenarios, however, is challenging to generalize to the full system where the complete cognition of the robot is formulated as a single POMDP in an end-to-end manner, because of the resulting high computational complexity. In reality, approximations of POMDPs are used to solve parts of the cognitive process, for instance, self-localization with a Particle Filter [9] or object tracking with a Multi-Hypothesis Kalman Filter [23].

Cognition in a complex robot is usually realized as a number of heterogeneous modules and services, which process and integrate sensory information into models of its environment. Such models might perhaps describe the position of the robot

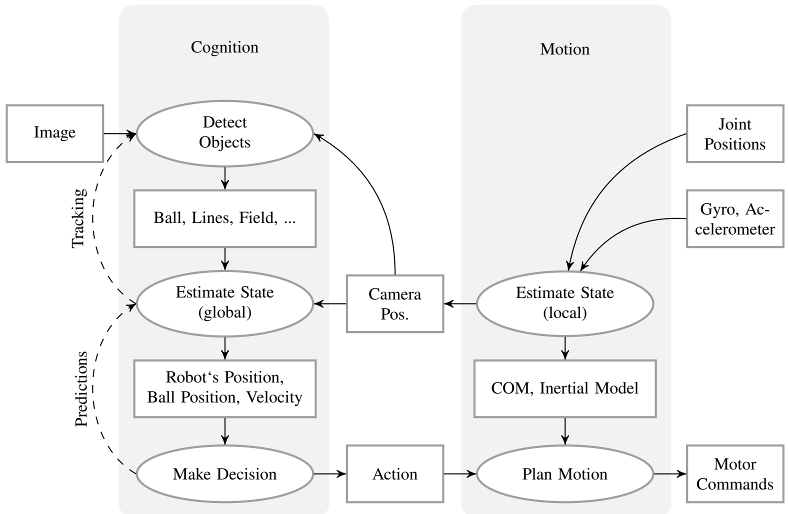

Fig. 1 Simplified example of a cognition in a humanoid soccer-playing robot. Round nodes depict modules and services responsible for processing information. Rectangular boxes depict data representations. The overall process is divided in two parts labeled **Cognition** and **Motion**. **Motion** contains critical sensorimotor functionality, e.g., keeping balance, and is executed with higher frequency with guaranteed execution times. **Cognition** groups modules responsible for higher cognitive functionality and is executed with lower frequency.

in its environment, positions of other objects, people, or robots. These models are updated over time and might be a result of different types of algorithms. These models, at a fixed time t, are also called *state* of the robot at the time t.

Figure 1 illustrates an example of a cognitive process implemented in a humanoid robot playing soccer in RoboCup. The process is divided in modules (round nodes) processing information, for instance, detecting the ball in the image and calculating its coordinates or estimating robot's position on the field. Each of the modules can be realized with a different approach, for instance, a deep neural network could be used for object detection, and a particle filter for the estimation of the robot's position on the field. The results, such as coordinates of a soccer ball detected in the image or position of the robot on the field, are stored and communicated between the modules through representations (rectangular boxes) The decision module, denoted by **Make Decision**, uses the estimated state of the situation to choose and to plan actions.

This subdivision provides high flexibility in choice of algorithms. At the same time, the decision mechanism, depicted by the node **Make Decision**, needs to be able to infer coherent decisions based on heterogeneous data produced by different modules. One of the common approaches to implement the decision mechanism are heuristic rule-based systems, such as widely used CABSL[1] [43] or XABSL[2] [29]. In

[1] `https://github.com/bhuman/CABSL` (accessed on 8.8.2022)

[2] `https://www.sim.informatik.tu-darmstadt.de/xabsl/index.html` (accessed on 8.8.2022)

such cases the goals of the robot are implicitly encoded in manually designed rules, which can lead to high complexity and errors.

The state estimation can be seen as an abstraction layer between a symbolic decision system and uncertain, noisy, ambiguous and incomplete sensory data. The quality of the decision depends on the quality of the state estimation, i.e., how well the estimations correlate with reality, e.g., precision of the estimated position of the robot. One could argue, that in these cases, the complexity and intelligence of the resulting behavior exhibited by the robot stems directly from the state estimation.

A system based on prediction and anticipation, on the other hand, allows to formulate the goals of the robot in an explicit and robust way.

3.2 Computational Anticipation

In this subsection, we consider practical aspects of the realization of the anticipatory decision process.

The basic principle of the anticipatory decision process can be summarized in a straightforward way: envision the possible futures resulting from possible actions and choose the action promising the best future.

From a practical standpoint, it is easy to see that an algorithm computing all possible outcomes for all possible actions has exponential growth. To make the process tractable, we need a mechanism to reduce the number of envisioned futures to a few relevant ones. In other words, we need a mechanism to direct the attention to relevant future scenarios.

In order to realize this we need a simulator able to predict future scenarios. When thinking about any kind of simulator (or a predictive model), two main questions arise: concerning the discretization of space and time. For instance, a recurrent neural network predicting a sensory measurement of a pressure sensor runs in fixed time steps, e.g., 10ms and produces numerical values in the working range of the sensor.

The situation becomes less obvious when we consider more complex scenarios on a higher decision level, e.g., a humanoid robot needs to make a decision regarding in which direction to kick the ball. Here, the question concerns the relevant parts of the environment to be simulated and how they to be represented. Evidently, we cannot simulate everything, so a reasonable simplification needs to be done here. For instance, some objects might be described by their position and their movement vectors. We also might limit the simulation only to the objects and agents involved in the situation.

Another important question concerns time. Typical physical simulators execute computations in small equidistant time steps to ensure realistic approximations of physical laws. This, again, might entail large computational effort. Another possibility for discretization in time is based on considering key-events. For instance, when the ball is kicked, the next interesting point in time could be when the ball comes to a standstill or enters the goal. Intermediate movement of the ball can be simplified, as long as influences relevant to the final outcome of the event are captured. This way, time can be split into distinct events and only the state of the situation in those events needs to be simulated.

The final issue that we discuss here is uncertainty. Any prediction in a real-world scenario will have some degree of uncertainty. We can identify several different sources of uncertainty. The first is the uncertainty coming from perception and state estimation, e.g., position of the robot on the field might be shifted, the same goes for the perceived objects. The second component stems from the uncertainty of the robot's actions. For instance, when the robot kicks the ball, the ball doesn't necessarily follow the same trajectory every time, the trajectory instead varies based on the exact location where the ball was hit, orientation of the grass blades, etc. The third source are the discretization and simplification artifacts introduced by the prediction process itself.

3.3 Computational Models of Anticipation

The intuition behind a simulation-based approach is to imagine (or simulate) what would happen as a result of execution of a particular action and then choose the action with the optimal (imagined/simulated) outcome.

In this subsection, we attempt a mathematical formulation for a decision process based on the principle of anticipation. The aim of this subsection is to ground and guide our intuition when talking about anticipation and decision-making. Formulating our ideas in a formal way will illuminate aspects otherwise not easily visible. The aim is not to arrive at a closed theory of anticipation, but rather to structure intuitive ideas and ground them in relation to each other. For our considerations, we will borrow a basic set of tools from function theory, set theory, probability theory and reinforcement learning.

Let's assume that the relevant state at the point in time t can be described by a vector $s_t \in \mathcal{S}$ with the state space $\mathcal{S}$. Let further $a_t \in \mathcal{A}$ describe a possible action at time t in the action space $\mathcal{A}$. In general, we can assume the state and the action to be real vectors with $\mathcal{S} \subseteq \mathbb{R}^n$ and $\mathcal{A} \subseteq \mathbb{R}^m$.

3.3.1 Forward Model.

An explicit predictive model can be written as a function

$$f : \mathcal{S} \times \mathcal{A} \to \mathcal{S} \tag{1}$$

$$(s_t, a_t) \mapsto f(s_t, a_t) := s_{t+1} \tag{2}$$

which calculates the envisioned state s_{t+1} based on the current state s_t and the assumed action a_t. The function f can also be called an *explicit forward model*.

More generally, the model can be formulated as an implicit relation φ between the states s_{t+1}, s_t and the action a_t. This relation can be written as

$$\varphi : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0,1] \tag{3}$$

$$(s_t, a_t, s_{t+1}) \mapsto \varphi(s_t, a_t, s_{t+1}) := p_t \tag{4}$$

The value of the function φ describes how likely the state s_{t+1} will occur in the future after the execution of the action a_t, given the current state s_t. We could call the function φ an *implicit forward model*.

The main advantage of an explicit model f is that the future state is directly calculated. This can be done, for instance, with forward simulation. Note, f is a deterministic model because it computes a single possible outcome. The implicit model φ can be more expressive. It can reflect the degree to which the future state is likely to occur and, more importantly, several future states can have equally-high likelihood of occurrence. In other words, φ is a probabilistic model. Model φ can be used to compute a likely outcome of an action in a *maximum likelihood* fashion

$$\hat{f}(s_t, a_t) := \operatorname*{argmax}_{s \in \mathcal{S}} \varphi(s_t, a_t, s) \tag{5}$$

This is, however, not necessarily deterministic, as several possible future states can have the same highest likelihood. In such cases, it will depend on the particular implementation of function argmax, which of the states is selected as the result.

Note, in probability theory, the expected value of s_{t+1} can also be calculated as an average of all possible states weighted with their probabilities. The value calculated this way is also called the *Bayes-hypothesis*. In our context, this approach only makes sense if the state space is continuous and the probability distribution φ is unimodal. Otherwise, in the worst case, it could happen, that the predicted state is very unlikely or might not even be possible, i.e., it might not be in the state space. An example could be a situation where a soccer-playing robot is directly facing a goal post with the ball placed directly in front of it. A forward kick would lead to a collision of the ball with the goal post and result in the ball being deflected to either side of the goal post with equal probability. If we had calculated the expected probabilistic state for this situation, then we would receive a position of the ball directly behind the goal post as an average position of the left and right positions, which is an impossible result.

In a way the forward model captures the geometry and physics of the situation in which the action is executed.

3.3.2 Inverse Model.

We previously defined forward models, which implicitly or explicitly predict the future state given the knowledge of the current state and an action. The decision problem could be formulated as an inverse of prediction. We want to decide which action to take in order to reach a particular desired state.

Let's consider the explicit forward model f. The current state $s_t \in \mathcal{S}$ is fixed and cannot be changed, thus it could be considered more of a parameter vector, as long as we are considering the fixed time point t and write

$$f_{s_t}(a) := f(s_t, a) \tag{6}$$

Let $s_{t+1} \in \mathcal{S}$ be the desired state. The task of finding an action $a_t \in \mathcal{A}$, that would result in the state s_{t+1} can be formulated as an inverse of f_{s_t}. Please note, that the function f_{s_t} is not necessarily invertible, since there might be several actions that lead to the same target state, i.e., $f_{s_t}(a_1) = f_{s_t}(a_2)$. Thus, the result of the inverse function f^{-1} is a set of all possible action, that would result in x_{t+1}, given the current state s_t.

$$a_t \in f_{s_t}^{-1}(x_{t+1}) := \{a \in \mathcal{A} | f_{s_t}(a) = x_{t+1}\} \tag{7}$$

With this, the decision algorithm computing an action to be executed is not deterministic in general. As a consequence, we need an additional mechanism to select one action. Such a selection can be done based on criteria specific to the action space $\mathcal{A}$. For instance, we could prefer a quicker action, or a less riskier one. In order to express to what degree a given action is desirable, we formulate a risk function

$$r : \mathcal{A} \rightarrow \mathbb{R}_+ \tag{8}$$

$$a \mapsto r(a) \tag{9}$$

The function captures the risks associated with a particular action a. In general, the risk function could be also dependent of the current state as well.

3.3.3 Value Function.

When making decisions in complex real-world scenarios, there is a set of possible desired states. For instance, in robot soccer it is desired to get the ball inside the opponents goal, the precise place inside the goal box is, however, irrelevant. This subset of goal states $S_G \subseteq \mathcal{S}$ might have a compacted shape within the state space $\mathcal{S}$. Additionally, the goal states might be not expressed exactly, but rather the desirability of states might be continuous, i.e., the closer to the opponent goal, the better. This can be captured in a *value function* describing desirability or *value* of every state

$$v : \mathcal{S} \rightarrow \mathbb{R}_+ \tag{10}$$

$$s \mapsto v(s) \tag{11}$$

The value function v encodes the value of a particular state with the goal of the systems to maximize the value.

3.3.4 Action Selection.

Making a decision implies selecting a single action from the set $\mathcal{A}$ to be executed. Previously, we formulated two criteria for the selection of an action, which are captured in the risk and value functions r and v. The aim of the decision algorithm is to minimize r and to maximize v over all possible future states. To formulate this, we can define a *utility function* combining the risk of the action and the value of the state.

$$u : \mathcal{S} \times \mathcal{A} \to \mathbb{R}_+ \tag{12}$$
$$(s, a) \mapsto u(s, a) := v(s) - r(a) \tag{13}$$

With this, the decision mechanism based on the explicit predictive model f can be formulated as a function

$$s : \mathcal{S} \to \mathcal{A} \tag{14}$$
$$s_t \mapsto a_t := s(s_t) := \operatorname*{argmax}_{a \in \mathcal{A}} u(f(s_t, a), a) \tag{15}$$

The selection function s assigns each state s_t the action a with the highest value $v(a)$. For s to be well defined we have to assume that the action space $\mathcal{A}$ is not empty. Essentially, function s searches for the action a_t that promises to lead to the future state s_{t+1} with the highest value $v(s_{t+1})$. Function s can be non-deterministic and can have more than one solution, as multiple future states can have the same highest value and different actions might lead to the same future state. In other words, the functions v and f are not necessarily invertible with respect to the argument a.

The state-action space $\mathcal{S} \times \mathcal{A}$ covers all possible combinations of states and actions $(s, a) \in \mathcal{S} \times \mathcal{A}$ and the function f predicts a future state for all combinations. In a real world scenario not all actions must necessary be possible or meaningful in all states. The function f describes what would happen if the action would be attempted. Depending on the concrete scenario the future state might remain the same, meaning $f(s, a) = s$, or result in state with lower value $v(s_{t+1})$ representing the failed attempt. For instance, picture a soccer robot lying face-down on the floor after a fall, in this state executing a kick is not possible and an attempt might be damaging to the robot. If the kick is attempted, then the situation would remain the same, or perhaps contain a malfunctioning knee joint and decreased energy level. In this scenario the kick action would have a significantly lover value compared to a *stand-up-motion* or *do-nothing* action.

Let's consider the combined function $Q := v \circ f$ evaluating state-action pairs

$$Q : \mathcal{S} \times \mathcal{A} \to \mathbb{R}_+ \tag{16}$$
$$(s_t, a_t) \mapsto r := (v \circ f)(s_t, a_t) = v(f(s_t, a_t)) \tag{17}$$

Using terminology from Reinforcement Learning (RL) we see that the function Q corresponds to the *state-action value function* used in Q-Learning [53]. Function Q predicts the expected value (reward) to be received when the action a_t is executed in the state s_t. The formulation in RL does not usually split function Q in the sub-components of prediction f and evaluation v. Instead, state-action function Q is formulated in a parameter-free manner, e.g., grid of neural network. This means that the ability to predict state transitions, and to evaluate the states are trained simultaneously. We see here that a basic formulation of the decision problem in RL implies anticipation as a basis for making decisions. One of the most recent studies on application of RL published in [28] shows that the ability to predict future states of the situation emerges in humanoid agents trained to play soccer in simulation. Moreover, this ability positively correlates with agent's performance.

Let's consider the implicit formulation of the decision algorithm.

$$s_i : \mathcal{S} \to \mathcal{A} \tag{18}$$

$$s_t \mapsto a_t := s_i(s_t) := \underset{a \in \mathcal{A}}{\operatorname{argmax}} \int_{\mathcal{S}} \varphi(s_t, a, s) \cdot v(s) \mathrm{d}s \tag{19}$$

Given action a and current state s_t, the function $\varphi(s_t, a, s)$ describes the likelihood of the future state s. The integral $\int_{\mathcal{S}} \varphi(s_t, a, s) \cdot v(s)$ describes the expected value of the action a. In case $\varphi(s_t, a, s)$ is a probability function this corresponds to the expected value

$$\mathbb{E}(a|s_t) = \int_{\mathcal{S}} \varphi(s_t, a, s) \cdot v(s) \mathrm{d}s \tag{20}$$

3.4 Calculating Anticipation

To make a decision in a concrete scenario, we need to compute values for the Equation (19). In general, this can pose a considerable challenge, as the models can take on complicated shapes making explicit calculation intractable.

One possible way to overcome this challenge is approximation through sampling. Sampling is widely used to implement various probabilistic methods in robotics, specifically for state estimation. Particle-Filters used for self-localization are one such examples. For an action $a \in \mathcal{A}$ and a fixed current state s_t we draw a fixed number N of samples for the possible future states with respect to the (density) function $\varphi(s_t, a, \cdot)$. The set of those samples comprises a *hypothesis* for the future state:

$$\mathcal{H}_a^N := \{s_i \in \mathcal{S} | s_i \sim \varphi(s_t, a, \cdot) \wedge 0 \leq i \leq N\} \tag{21}$$

4 Decision-Making in a Bio-Hybrid Beehive

A bio-hybrid beehive is a beehive equipped with additional sensors, actuators and computational resources, which allow the bee colony to better cope with adverse factors in challenging environments and eventually reduce competition for resources between honeybee colonies and other bee species. This augmented beehive can be considered a robot able to make autonomous decisions and act on them.

In this section, we discuss how the principle of anticipation can be applied for decision-making in a scenario involving bio-hybrid beehives. We build and expand on our preliminary work published in [51], where we introduced a heuristic rule-based process for active selection of foraging sources in a bio-hybrid beehive. The approach was implemented and tested using an agent-based simulation for the foraging behavior of the bees.

In the following, we first introduce bio-hybrid beehives. Afterwards, we present an abstract simplified simulation for the behavior of bees, which can be used as an internal simulation in an anticipatory decision-process of a bio-hybrid beehive. In the third part we discuss example scenarios and formulate the decision process. We close the section with remarks on experiments and implementation.

4.1 Intelligent Bio-Hybrid Beehives as Autonomous Robots

In this section, we provide a brief introduction in specific areas of research of honeybees to provide grounding for our later investigations. Specifically, we illuminate those studies which can be a basis for a connection between the discipline of robotics and studies of bees.

There is a large variety of bee species and some of them have been successfully domesticated by humans. We will refer to the domesticated bees as *honeybees*. Honeybees are primarily used in two major areas: honey production and pollination services in agriculture. In the latter, beehives are intentionally placed in proximity to agricultural fields that require pollination in order to increase yield. Extensive proliferation of honeybees can make them into an invasive species, displacing wild bee species in competition for limited habitable space, as well as foraging resources, such as nectar and pollen. Examples for studies investigating the role of honeybees in extirpation of wild bees can be found in [21] and more recently in [42].

The European Project HIVEOPOLIS [10,22] aims at developing a true bio-hybrid symbiotic system. The bio-hybrid beehive will be equipped with a wide range of sensors and processing power to monitor and support the health of the colony.

Additionally, the beehive will be equipped with an internal robotic actuator interacting with the bees in the colony through an imitation of a *waggle dance*. In a bee colony, successful foragers can share information about the foraging locations through a specific *waggle dance* [17,18]. A prototype of a robot imitating the waggle dance, called *RoboBee*, was introduced in [26]. The robot is installed in the beehive and moves a dummy-bee imitating waggle dance, similar to a real bee. It has been shown, that *RoboBee* was able to encourage a significant number of bees to fly to a specific foraging location. The study in [27] investigates the expected effect of dancing robots on the behavior of the bees in a simulated scenario. On the other hand, a waggle dance performed by the bees can be suppressed to prevent them from flying to an undesired location, for instance, because of the contamination with pesticides. Such suppression can be achieved using vibration actuators embedded in beehive as discussed in [50]. A bio-hybrid beehive is connected to a network that allows sharing information with other bio-hybrid units and provides access to a centralized database with extended information regarding possible flowering locations,

locations of other beehives in the neighborhood, protected areas reserved for wild bees, weather conditions, and further relevant environmental information.

With this in mind, a bio-hybrid beehive can be seen as a *robot* able to make autonomous decisions and negotiate with other beehives, e.g., send the bees to a particular region or, potentially, prevent them from harvesting in another.

4.2 Simulation

At the core of the decision mechanism is the predictive model which allows to predict the outcome of the beehive's actions. This predictive model can be realized as an internal simulation. In this section, we discuss a simplified agent-based simulation for the behavior of bee colonies that can be used to develop and study decision mechanisms for bio-hybrid beehives.

Bees are able to exhibit complex swarm behaviors like decentralized target selection and workload balancing. The behavior of bees has been extensively studied. Specifically, a wide variety of simulations have been proposed, which allow to synthetically investigate behavioral dynamics of the bee swarms. Schmickl and Crailsheim introduce in [48] a multi-agent simulation, which is able to simulate the dynamics of honeybee nectar foraging. The authors implemented experiments reported in [49] and other works of T. Seeley, who investigated natural decision-making mechanisms within a bee swarm. Multi-agent approach was widely employed to study foraging behavior of the bees, including works by Dornhaus and colleges [14], and Beekman and colleges [4]. Well known models BEESCOUT [2] and BEEHAVE [3] are used for better understanding and exploration of the possible realistic scenarios of natural colony dynamics, bees' searching behavior in habitats with different landscape configuration, as well as interactions between bees within a colony. Another example for an agent-based simulator for honeybee colonies was published in [5]. Finally, in [27], authors investigate the effect of robotic actuator imitating bee dance on the bee colony's foraging decisions in a bio-hybrid beehive using mathematical models.

Although the above models were created from the perspective of studying the behavior of bees, they can be used in an internal simulation of an intelligent bio-hybrid beehive to predict the effects of its actions and the actions of other colonies in its surroundings in order to make decisions on the swarm level based on those predictions.

We investigate how a decision mechanism based on anticipation, as discussed in Section 3, can be realized with the help of such internal simulation. For this, we implement a simplified simulation for the behavior of bee swarms capturing only the essential aspects. For this, we extend on our preliminary work [51], where we presented an agent-based simulation for foraging behavior of bees.

We implement the simulation as a multi-agent system, since this approach has proven successful in simulating the behavior of beehives. The simulation is implemented as a multi-agent system with three basic types of agents: *bee*, *beehive*, and

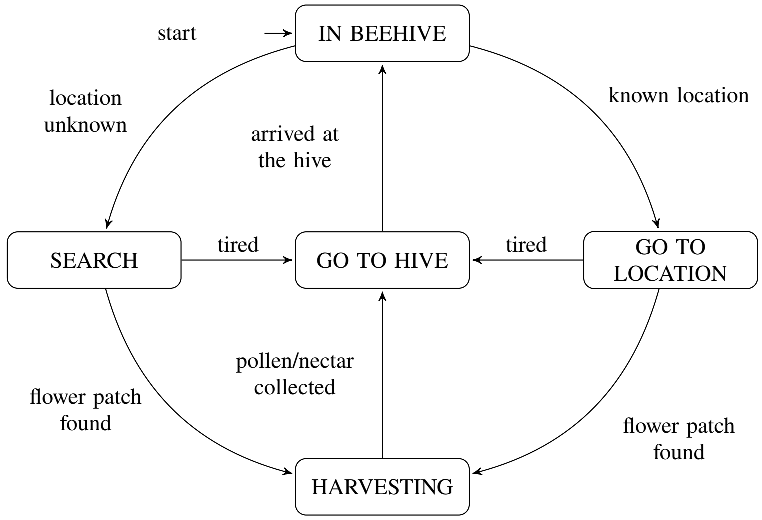

Fig. 2 State diagram describing the behavior of a bee agent.

field. The environment is modeled as a discrete squared grid. Each cell of the grid contains an agent *field* and can additionally contain other agents, which can be an agent *beehive* and a number of *bee* agents. All *bee* agents in the same cell as a *beehive* are seen as being inside the *beehive*. The same applies for the *field* – all *bee* agents in the same cell are considered to be foraging in that field. All agents know the coordinates of their cell and have access to the list of other agents in the same cell. Each agent has a location on the grid, i.e., coordinates of the cell in which it is located, a number of values describing its internal state, and an activation function, which is called in each step of the simulation. The simulation and the agents are updated in discrete time steps. The activation order of agents is randomized in order to reduce its impact on the model.

The agents *field* and *beehive* represent places with which bees can interact; their own functionality is limited. The agent representing a foraging *field* is implemented in a simple way. The *field* can be in a blooming state, which can change depending on time to simulate blooming periods. The *field* has also a value representing available resources, which can be harvested when the field is blooming. The agent representing *beehive* is mainly responsible for holding information about the state of the colony, like the amount of collected honey. Other than that, the *beehive* does not have its own functionality. It represents a regular beehive without any technical extensions.

The agent *bee* models the foraging behavior of a honeybee. The behavior is implemented as a state-machine shown in Figure 2. We differentiate between five states: `IN BEEHIVE`, `SEARCH`, `GO TO LOCATION`, `HARVESTING`, and `GO TO HIVE`. In each state a bee executes a number of actions specific to that state and decides whether to change to another state in the next time step. The bee stays in the current state until conditions for a transition to another state are met. In Figure 2, the self-transitions are not visualized for simplicity. The internal state of a bee is described mainly by the energy that a bee has, a value representing the amount of resources it is carrying, and the current state of the behavior, as described above.

`IN BEEHIVE` is the starting state. At the beginning of each simulation all bees are located in their beehives. When inside the beehive, a bee performs three tasks: transferring collected pollen to the hive, recharging energy, and recruiting other bees to its foraging location if foraging was successful. Which bees are recruited and whether recruiting is successful is decided probabilistically, simulating a waggle-dance.

`SEARCH` is executed if a bee leaves the hive without a known foraging location. In each step one of the neighboring cells is randomly selected as the next target to move to. The location from the past step is excluded to prevent oscillations between two cells. With each move the amount of energy is reduced.

`GO TO LOCATION` works similar to `SEARCH`, but is more focused on a specific target location. If a bee has a known foraging location, the neighboring cells closer to the target are chosen with higher probability. This means that the bee is gradually pulled towards its target location and can still explore the environment along the route.

`HARVESTING` – If the current cell contains a blooming *field*, then transition to `HARVESTING` is made. In the state `HARVESTING` the bee collects pollen and nectar if available and leaves after a fixed length of time by transitioning into the state `GO TO HIVE`.

`GO TO HIVE` is implemented in the same way as `GO TO LOCATION`. In this case the target is the location of the own hive. A bee transitions into `GO TO HIVE` after `HARVESTING`, and also from the states `SEARCH` and `GO TO LOCATION` in case the bee gets tired and runs low on energy.

Figure 3 illustrates an example run of the simulation in a simple scenario with a single beehive and one field with resources. The plot at the bottom in Figure 3 shows the amount of resources (pollen, nectar) collected by bees at each step of the simulation. The collected amount of pollen is represented in an abstract measurement unit. At the beginning of the simulation we see the bees exhibiting the random search behavior. At the time $t = 105$, one of the bees returns with nectar and begins communicating the coordinates of the field to the others. From that time onward, we can observe the behavior of the bees gradually changing to directed harvesting from the field, as the information about its location is propagated among the bees. The change in behavior can also be observed on the shading of the cells, which reflects the number of times a cell was visited by the bees.

To simulate the bio-hybrid beehive, we extend the simulation by adding two new agents *RoboBee* and *BioHybrid*. *RoboBee* is an agent that always stays inside its beehive and participates in recruiting other bees for a certain foraging location. The only task of a *RoboBee* is to hold a provided target location for foraging and to participate the in recruiting of other bees. *BioHybrid* is an extension of the agent *beehive*; it can contain several *RoboBee* agents. Additionally to a *beehive*, *BioHybrid* executes a non-trivial functionality, when activated during the simulation. In each time step, *BioHybrid* has the ability to decide on target foraging locations, which

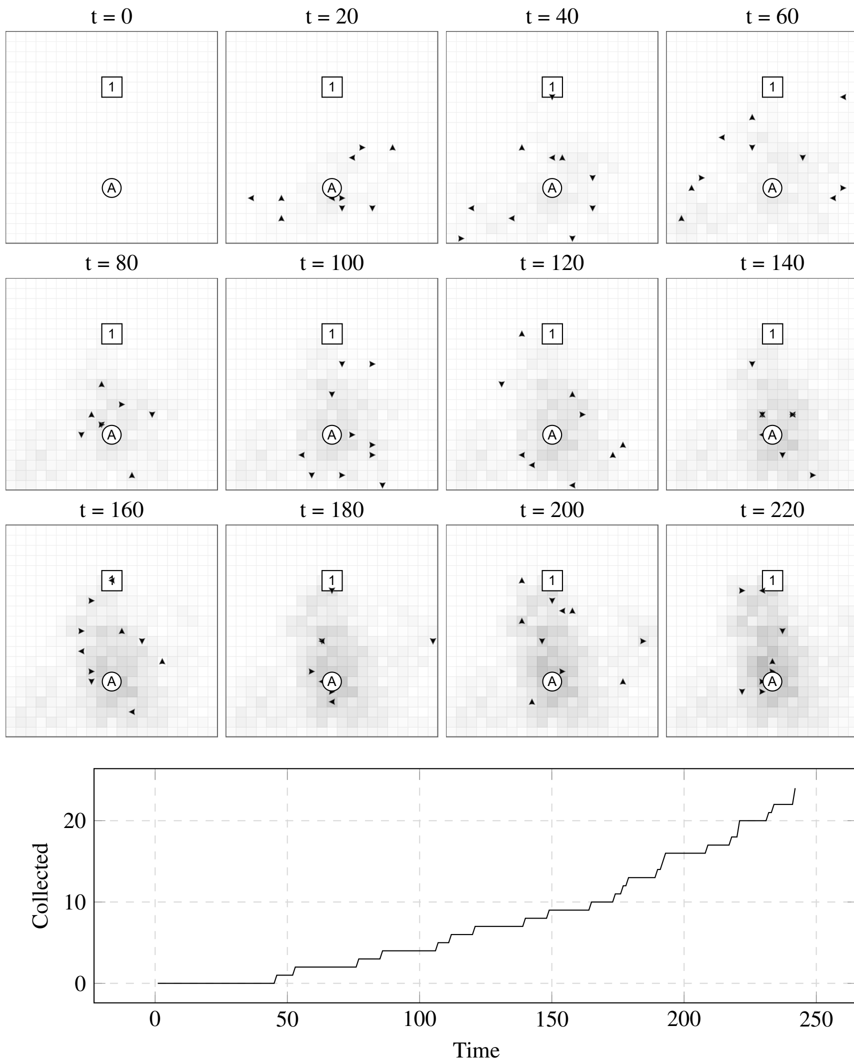

Fig. 3 Example for a simulation run with one beehive (circle) and one foraging ground (square). The bees are depicted by the black arrows. The snapshots of the simulated environment are taken every 20 time steps. The shading of the cell indicates how often this cell was visited by a bee. The graph below shows the amount of collected resources (pollen, nectar).

are communicated to its *RoboBee*s. Figure 5 illustrates an example for an active selection of foraging locations by a *BioHybrid* equipped with one *RoboBee*. In the left part of Figure 5, no active selection is done; the *RoboBee* is deactivated and the bees behave in the same way as an uncontrolled (regular) beehive. On the right side of the Figure 5, the *BioHybrid* actively selected the fixed location of the field further away from the hive.

We focused on the simulation of interactions between beehives as agents, therefore, bees' complex behavior was reduced to foraging behavior. For the sake of simplicity, we do not model the complex social organisation of an individual colony, ignoring the diversity of bees' casts (workers, drones, queen) and food sources (nectar, pollen, water). We implemented two foraging strategies for bees agents: *random foraging search* and *targeted foraging on a known patch of flowers*. In the first case, bees randomly explore the environment around the hive in search of a flower field. In the second case, bees have knowledge about the location of a flower field and head directly towards it for foraging. The simulation has shown promising results for being suitable for studying the decision-mechanism for bio-hybrid beehives.

4.3 Anticipation and Decision-Making for Beehives

In this section, we discuss how anticipation can be used to realize decision-making in a bio-hybrid beehive. We want to enable the bio-hybrid beehive to act autonomously within the ecosystem in a way that is beneficial to all involved. The ability to direct bees to certain areas while avoiding others could lead to a wide range of scenarios benefiting all participants of the symbiotic relationship: humans (consumers, beekeepers, farmers), honeybees and wild bee species. For instance, sending bees to a known foraging ground with a high yield of nectar and pollen or plants of a certain kind could lead to higher honey harvest and ensure a specific type of honey. Sending bees to certain locations, which require pollination could increase the quality of the pollination service provided by the bees. Avoiding areas known to be contaminated with pesticides could contribute to bees' health and well-being. Habitats and natural foraging grounds of wild bee species could be protected by avoiding reserved areas.

The bio-hybrid beehive employs a robotic actuator inside the beehive to imitate the waggle dance of the bees and to *recruit* the bees to forage at a certain location. The resulting behavior of the bees is highly complex, since the bees behave autonomously and not all bees necessarily follow the robotic actuator. The actual flight paths of the bees depend highly on the environment and choices of the bees. This means that the actions of a bio-hybrid beehive have complex consequences with a high degree of uncertainty. To make an effective decision in these circumstances, a decision-mechanism will require the knowledge of the environment and understanding of the interactions between the bees the environment and other bee-species. In our previous work [51] we investigated a heuristic rule-based approach, which was shown to work in certain scenarios, but also quickly became difficult to maintain and error-prone due to complexity of the system. To manage the complexity in the behavior of the bees, we investigate a decision-making mechanism based on anticipation as discussed in Section 3.

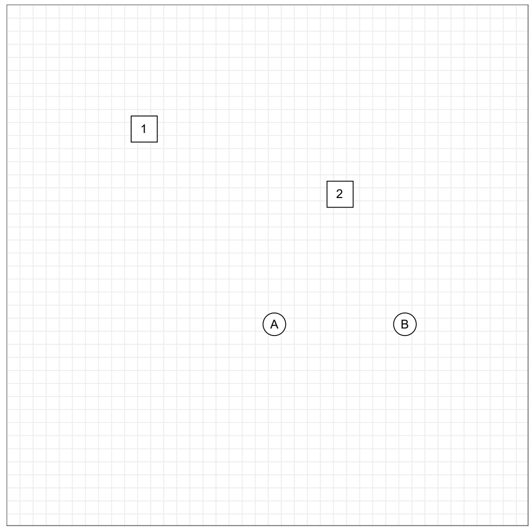

Fig. 4 Example scenario: Field 1 is located at the same distance to bio-hybrid beehive A and the wild beehive B. Field 2 is out of reach for wild beehive B, but can be reached by the bio-hybrid A.

We begin with a simple scenario as illustrated in Figure 4. The scenario consists of two beehives – a bio-hybrid beehive **(A)** and a wild beehive **(B)**, and two fields **(1)** and **(2)** located in the proximity of the beehives, where bees can forage nectar and pollen. The foraging ground **(2)** is located closer to both beehives, while the other **(1)** is out of reach for the bees from the wild beehive, but still reachable by the bees from the bio-hybrid.

As one of the flower fields is closer to both beehives, it is most likely to be the preferred foraging ground for both colonies. This can lead to a competition for the same resource between the beehives and eventually cause a shortage of food for the wild bee colony. As a result, the wild bee colony might end up unprepared for the next winter period or being forced to migrate to another location. Figure 5 (left) illustrates this scenario.

In the second scenario, we assume that the bio-hybrid beehive is aware of the location of the wild beehive as well as other flowering fields in the neighborhood. In order to protect the wild bees, the bio-hybrid beehive could actively encourage its bees to forage at the field **(1)** located at a larger distance. Despite being less attractive, the suggested alternative field still provides enough resources. This approach may decrease the amount of foraged pollen and nectar for the bees of the bio-hybrid beehive because of the distance to the foraging field. Nevertheless, the overall yield of both beehives would still be sufficient and the wild beehive would be protected.

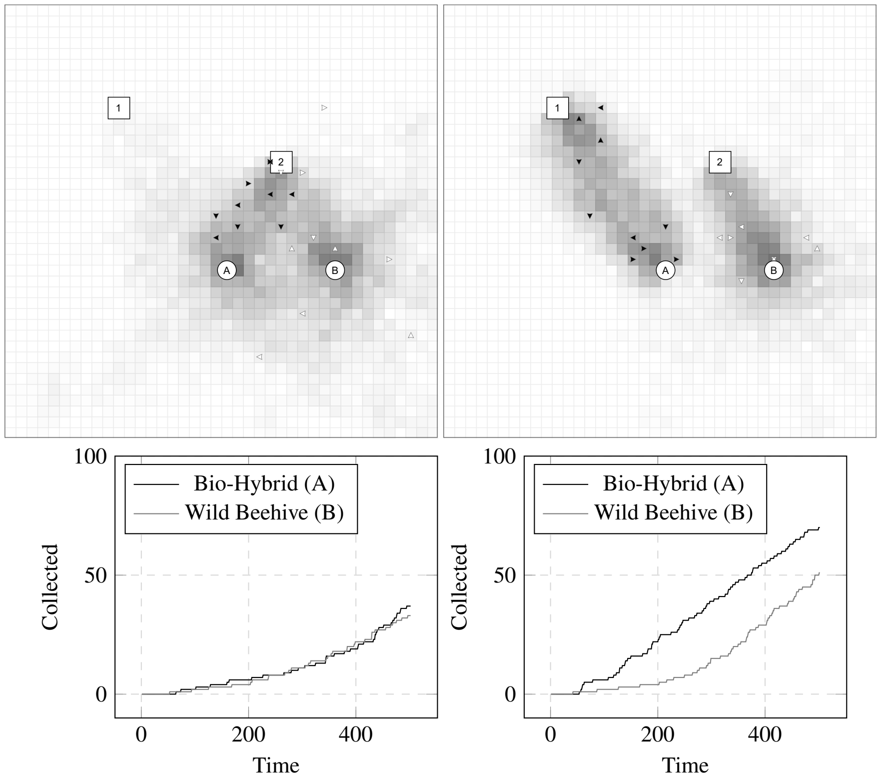

Fig. 5 Two scenarios for decision making in bio-hybrid beehives after 500 simulation steps. Bio-hybrid beehive A, a wild beehive B are depicted by circles, and two foraging fields 1 and 2 by squares. Both fields have 100 resources (nectar, pollen). Shading of the cells indicates the number of times the cell was visited by a bee. The number of resources collected by each colony in both scenarios is illustrated by the corresponding plots. Left: with no active control, both beehives A and B choose the closest flower patch 2 and compete for the resources. Right: the bio-hybrid beehive A actively motivates its bees to choose a less convenient foraging location 1 to give the wild beehive space and to improve overall yield.

In this case, the bio-hybrid beehive collaborates with the wild beehive instead of competing for the same resources. This scenario is illustrated in Figure 5 (right).

In order to make the decision to choose a less convenient option for the sake of common benefit, the bio-hybrid beehive must be aware of the wild beehive's actions. Direct observation of the behavior of the wild bees cannot be realized in a practical way. However, we can use simulation to predict the behavior of the wild bees, assuming we know the location(s) of their beehive(s), distribution of the flowering fields in the neighborhood, as well as a model for their behavior.

We formulate the decision problem more generally. *BioHybrid* could predict the amount of resources collected by each hive for each of the possible foraging locations it could select. Based on that, the target field with the most fair distribution

Algorithm 1: Action selection based on internal simulation

Data: S_0, $foraging\_fields$
Result: a^*, r^*
$\mathcal{A} \leftarrow foraging\_fields \cup \{none\}$
`// simulate: run the simulation for each action` a
$consequences \leftarrow \{\}$
for $a \in \mathcal{A}$ **do**
    $s_0 \leftarrow S_0$
    for $t \in [0, T]$ **do**
        $s_{t+1} \leftarrow simulate(s_t, a)$
    end
    $consequences \leftarrow consequences \cup \{(a, s_T)\}$
end
`// evaluate: minimal collected ressources by a beehive`
$results \leftarrow \{\}$
for $(a, s_T) \in consequences$ **do**
    $r \leftarrow \min(collected\_ressources(s_T))$
    $results \leftarrow results \cup \{(a, r)\}$
end
`// select: action with maximal predicted value`
$a^* \leftarrow none$
$r^* \leftarrow 0$
for $(a, r) \in results$ **do**
    if $r \geq r^*$ **then**
        $a^* \leftarrow a$
        $r^* \leftarrow r$
    end
end

of projected harvest can be selected. As formulated in Section 3 we need to define components: predictive model, evaluation function, and a selection mechanism.

An *action* executed by the agent *BioHybrid* consists of a target foraging location communicated to the bees in the hive through the *RoboBee*. It is also possible for *BioHybrid* to deactivate *RoboBee*, such that no target foraging location is communicated and the bees are left to forage on their own device. We could call this action to be *neutral*. In this scenario, the set of actions $\mathcal{A}$ consists of all reachable locations in the proximity of the beehive where the bees can be sent.

We consider a time period $[0, T]$ of a fixed length $T \in \mathbb{N}$ and assume that the decision is made only at the beginning of the simulated period, and the action selected by the *BioHybrid* agent remains the same for the whole time. The state s_t is precisely the state of the simulation at the time t. The value function $v(s_t)$ could be formulated as the minimal amount of honey collected by each beehive. That means that the aim is to maximize the least amount of honey collected by a hive, which, in an ideal scenario, would lead to all beehives collecting the same amount of honey.

The simulation presented above can be used as an internal simulator to predict the amount of collected honey for all beehives. Algorithm 1 outlines the three phases of the decision-mechanism: prediction (through simulation), evaluation and selection.

This simple scenario gives an illustration of a situation in which an intelligent bio-hybrid beehive can adapt its behavior to protect a wild beehive, which a regular hive of honeybees would not do. Of course, in reality a bio-hybrid beehive will be confronted with much more complicated scenarios.

4.4 Discussion

We implemented two scenarios described in the previous section. In the case of the first scenario, bees from both beehives search for flowering fields through random exploration of their neighbourhoods. As soon as a flowering field has been found, the information about it is communicated to the other bees (in reality it is achieved through waggle dancing). In case of the second scenario, the bio-hybrid beehive has information about food sources in the neighborhood which lie within the maximum flying distance of bees from both beehives. It can estimate at which flowering field competition might eventually happen. If there is such field in the neighbourhood then it searches for the non-competitive fields, estimate which of the non-competitive fields might be the most attractive to the bees, and send them there. In case if all known floral resources are potentially competitive, then bio-hybrid beehive predicts for each of them the maximal number of competitive beehives and suggests one with lowest number of competitors. Of course, the second scenario is an ideal scenario. We should not forget that in reality bees are autonomous agents. They may follow the advertised information and fly to the suggested field but may also ignore it and continue foraging at the same field where the wild bees forage.

The experimental results show that simulation of different strategies can provide data for further analysis in order to define more precise parameters as well as fine tuning of the whole approach.

4.5 Remarks on Implementation

Simulations were implemented with the framework Mesa[3] [24]. In comparison to the other well-known simulation tools, like NetLogo[4] [52], Repast[5] [37] and Mason[6] [30], this framework has several competitive advantages. First of all, it is python-based and can be extended with modern python libraries and other python-based tools (e.g., Jupyter Notebook and Pandas tools) in order to create more complex simulations or analyse collected data. The collected data can be stored in a JSON or Pandas DataFrame format for further analysis. Second, Mesa consists of decoupled components, which can be replaced or used independently from each other.

[3] `https://mesa.readthedocs.io/` (accessed 08.08.2022)

[4] `https://ccl.northwestern.edu/netlogo/` (accessed 08.08.2022)

[5] `https://repast.github.io/` (accessed 08.08.2022)

[6] `https://cs.gmu.edu/~eclab/projects/mason/` (accessed 08.08.2022)

Third, visualization is browser-based, which provides additional opportunities for sharing of visualisation via the Internet. Since all components in the Mesa framework are decoupled, visualisation modules can be customized, extended, replaced, or removed.

5 Conclusions

Anticipation is a general principle used by humans and other animals to realize complex behaviors. In artificial agents, anticipation can be used to realize a powerful mechanism for inference and decision-making in complex scenarios.

We reviewed, how anticipation is studied in artificial agents and discussed possible formal foundations to ground the intuitive understanding of anticipation in case of decision-making. We discussed the scenario of a bio-hybrid beehive – a beehive equipped with robotic actuators allowing it to interact with bees and to encourage them to forage at certain locations. We implemented a simplified abstract simulation for the behavior of a bio-hybrid beehive. Results of the actions of a bio-hybrid have high complexity due to the high number of involved interactions between the bees and the environment. With the help of the simulation, the complexity can be reduced. Our preliminary experiments show that a decision made by a bio-hybrid beehive can help to protect areas of wild bees. This simulation forms a basis for further studies on a generalized decision mechanism for the bio-hybrid beehive.

To realize the complete decision mechanism for the bio-hybrid beehive we need three components: a predictive model (internal simulation), an evaluation function and a mechanism to select an action based on the value of its predicted outcome.

In its current form the simulator presents a proof-of concept. This is sufficient to study the basic principles of the decision process. In future work we will extend the simulation to better reflect real behavior of the bees and to bring it closer to an application in a real world scenario.

The evaluation function decides which future scenarios are more desirable and thus directly determines the behavior of the agent. For example, the amount of own collected honey, used as a value, would lead to egoistic competitive behavior. The example scenario shown in Figure 5 has demonstrated that it is possible to realize collaborative behavior by considering the envisioned amount of collected honey by others. This opens a question: how can we formulate an evaluation function encouraging collaborative behavior in a more general way? For this we plan to study collaborative and competitive behavior of the beehive in more complex scenarios involving a larger number of bio-hybrid and wild beehives.

Finally, we need to consider the selection mechanism. Because of the high computational complexity, it might not be possible to evaluate all available actions. Possible approaches to solve this could involve sampling.

In conclusion, anticipation in both animals and artificial system is a powerful mechanism for inference and decision-making in complex scenarios and could be studied for a variety of applications.

Acknowledgements This work was supported by the European Union's Horizon 2020 (H2020) research and innovation program under grant agreement No. 824069 (HIVEOPOLIS) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2002/1 "Science of Intelligence" – project number 390523135.

The authors would like to thank Vanessa Gorsuch for constructive criticism of the manuscript, which significantly contributed to readability and clarity of this publication.

References

1. Ahmadi, M., Stone, P.: RoboCup 2007: Robot Soccer World Cup XI, chap. Instance-Based Action Models for Fast Action Planning, pp. 1–16. Springer Berlin Heidelberg, Berlin, Heidelberg (2008). DOI 10.1007/978-3-540-68847-1_1. URL `http://dx.doi.org/10.1007/978-3-540-68847-1_1`
2. Becher, M.A., Grimm, V., Knapp, J., Horn, J., Twiston-Davies, G., Osborne, J.L.: Beescout: A model of bee scouting behaviour and a software tool for characterizing nectar/pollen landscapes for beehave. Ecological Modelling **340**(1), 126–133 (2016). DOI 10.1016/j.ecolmodel.2016.09.013
3. Becher, M.A., Thorbek, V.G.P., Horn, J., Kennedy, P.J., Osborne, J.L.: Beehave: a systems model of honeybee colony dynamics and foraging to explore multifactorial causes of colony failure. Journal of Applied Ecology **51**(2), 470–482 (2014). DOI 10.1111/1365-2664.12222
4. Beekman, M., Lew, J.B.: Foraging in honeybees—when does it pay to dance? Behavioral Ecology **19**(2), 255–261 (2007). DOI 10.1093/beheco/arm117. URL `https://doi.org/10.1093/beheco/arm117`
5. Betti, M., LeClair, J., Wahl, L.M., Zamir, M.: Bee++: An object-oriented, agent-based simulator for honey bee colonies. Insects **8**(1), 31 (2017). DOI 10.3390/insects8010031. URL `https://pubmed.ncbi.nlm.nih.gov/28287445`. 28287445[pmid]
6. Blum, C., Winfield, A.F.T., Hafner, V.V.: Simulation-based internal models for safer robots. Frontiers in Robotics and AI **4**, 74 (2018). DOI 10.3389/frobt.2017.00074. URL `https://www.frontiersin.org/article/10.3389/frobt.2017.00074`
7. Bongard, J., Zykov, V., Lipson, H.: Resilient machines through continuous self-modeling. Science **314**(5802), 1118–1121 (2006). DOI 10.1126/science.1133687. URL `https://science.sciencemag.org/content/314/5802/1118`
8. Bordallo, A., Previtali, F., Nardelli, N., Ramamoorthy, S.: Counterfactual reasoning about intent for interactive navigation in dynamic environments. In: Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on, pp. 2943–2950 (2015). DOI 10.1109/IROS.2015.7353783
9. Burchardt, A., Laue, T., Röfer, T.: Optimizing particle filter parameters for self-localization. In: J. Ruiz-del Solar, E. Chown, P.G. Plöger (eds.) RoboCup 2010: Robot Soccer World Cup XIV, pp. 145–156. Springer Berlin Heidelberg, Berlin, Heidelberg (2011)
10. CORDIS: Futuristic beehives for a smart metropolis (2019). URL `https://cordis.europa.eu/project/id/824069`
11. Demiris, Y., Khadhouri, B.: Hierarchical, attentive, multiple models for execution and recognition (hammer). Robot. Auton. Syst. J. **54** (2005)
12. Dermy, O., Charpillet, F., Ivaldi, S.: Multi-modal intention prediction with probabilistic movement primitives. In: F. Ficuciello, F. Ruggiero, A. Finzi (eds.) Human Friendly Robotics, pp. 181–196. Springer International Publishing, Cham (2019)
13. Dodds, R., Vallejos, P., Ruiz-del Solar, J.: Probabilistic kick selection in robot soccer. In: Robotics Symposium, 2006. LARS '06. IEEE 3rd Latin American, pp. 137–140 (2006). DOI 10.1109/LARS.2006.334337
14. Dornhaus, A., Klügl, F., Oechslein, C., Puppe, F., Chittka, L.: Benefits of recruitment in honey bees: effects of ecology and colony size in an individual-based model. Behavioral Ecology

17(3), 336–344 (2006). DOI 10.1093/beheco/arj036. URL `https://doi.org/10.1093/beheco/arj036`

15. Duarte, N.F., Tasevski, J., Coco, M.I., Rakovic, M., Santos-Victor, J.: Action anticipation: Reading the intentions of humans and robots. CoRR **abs/1802.02788** (2018). URL `http://arxiv.org/abs/1802.02788`
16. Eppner, C., Martín-Martín, R., Brock, O.: Physics-based selection of actions that maximize motion for interactive perception. In: RSS workshop: Revisiting Contact - Turning a problem into a solution (2017). URL `http://www.robotics.tu-berlin.de/fileadmin/fg170/Publikationen_pdf/EppnerMartinMartin17_RSS_WS.pdf`
17. von Frisch, K.: Tanzsprache und Orientierung der Bienen. Springer Berlin, Heidelberg, Berlin, Heidelberg (1965). DOI 10.1007/978-3-642-94916-6
18. von Frisch, K.: The Dance Language and Orientation of Bees. Harvard University Press, Cambridge, Massachusetts (1967). DOI 10.1007/978-3-642-94916-6. A translation of Tanzsprache und Orientierung der Bienen
19. Guerrero, P., Ruiz-del Solar, J., Díaz, G.: Probabilistic Decision Making in Robot Soccer, pp. 29–40. Springer Berlin Heidelberg, Berlin, Heidelberg (2008). DOI 10.1007/978-3-540-68847-1_3. URL `https://doi.org/10.1007/978-3-540-68847-1_3`
20. Hafner, V.V., Kaplan, F.: Interpersonal maps and the body correspondence problem. In: Proceedings of the Third International Symposium on Imitation in animals and artifacts, pp. 48–53. University of Hertfordshire, Hatfield, UK (2005)
21. Hudewenz, A., Klein, A.M.: Competition between honey bees and wild bees and the role of nesting resources in a nature reserve. Journal of Insect Conservation **17**(1), 1275–1283 (2013). DOI 10.1007/s10841-013-9609-1
22. Ilgün, A., Angelov, K., Stefanec, M., Schönwetter-Fuchs, S., Stokanic, V., Vollmann, J., Hofstadler, D.N., Kärcher, M.H., Mellmann, H., Taliaronak, V., Kviesis, A., Komasilovs, V., Becher, M.A., Szopek, M., Dormagen, D.M., Barmak, R., Bairaktarov, E., Broisin, M., Thenius, R., Mills, R., Nicolis, S.C., Campo, A., Zacepins, A., Petrov, S., Deneubourg, J.L., Mondada, F., Landgraf, T., Hafner, V.V., Schmickl, T.: Bio-Hybrid Systems for Ecosystem Level Effects. In: Proceedings of the ALIFE 2021: The 2021 Conference on Artificial Life, *ALIFE 2021: The 2021 Conference on Artificial Life*, vol. ALIFE 2021: The 2021 Conference on Artificial Life. MIT Press Direct, Virtual (formerly Prague), Czech Republic" (2021). DOI 10.1162/isal_a_00396. URL `https://doi.org/10.1162/isal_a_00396`. 41
23. Jochmann, G., Kerner, S., Tasse, S., Urbann, O.: Efficient multi-hypotheses unscented kalman filtering for robust localization. In: T. Röfer, N.M. Mayer, J. Savage, U. Saranlı (eds.) RoboCup 2011: Robot Soccer World Cup XV, pp. 222–233. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
24. Kazil, J., Masad, D., Crooks, A.: Utilizing python for agent-based modeling: The mesa framework. In: Social, Cultural, and Behavioral Modeling: 13th International Conference, SBP-BRiMS 2020, Washington, DC, USA, October 18–21, 2020, Proceedings, p. 308–317. Springer-Verlag, Berlin, Heidelberg (2020). DOI 10.1007/978-3-030-61255-9_30
25. Kunze, L., Beetz, M.: Envisioning the qualitative effects of robot manipulation actions using simulation-based projections. Artificial Intelligence **247**, 352–380 (2017). DOI https://doi.org/10.1016/j.artint.2014.12.004. URL `https://www.sciencedirect.com/science/article/pii/S0004370214001544`. Special Issue on AI and Robotics
26. Landgraf, T., Bierbach, D., Kirbach, A., Cusing, R., Oertel, M., Lehmann, K., Greggers, U., Menzel, R., Rojas, R.: Dancing honey bee robot elicits dance-following and recruits foragers (2018). URL `https://arxiv.org/abs/1803.07126`
27. Lazic, D., Schmickl, T.: Can robots inform a honeybee colony's foraging decision-making? In: Proceedings of the ALIFE 2021: The 2021 Conference on Artificial Life. MIT Press Direct, Virtual (formerly Prague), Czech Republic (2021). DOI 10.1162/isal_a_00397
28. Liu, S., Lever, G., Wang, Z., Merel, J., Eslami, S.M.A., Hennes, D., Czarnecki, W.M., Tassa, Y., Omidshafiei, S., Abdolmaleki, A., Siegel, N.Y., Hasenclever, L., Marris, L., Tunyasuvunakool, S., Song, H.F., Wulfmeier, M., Muller, P., Haarnoja, T., Tracey, B.D., Tuyls, K., Graepel, T., Heess, N.: From motor control to team play in simulated humanoid football. CoRR **abs/2105.12196** (2021). URL `https://arxiv.org/abs/2105.12196`

29. Loetzsch, M., Risler, M., Jungel, M.: Xabsl - a pragmatic approach to behavior engineering. In: 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5124–5129 (2006). DOI 10.1109/IROS.2006.282605
30. Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K.M., Balan, G.: Mason: A multi-agent simulation environment. SIMULATION **81**(7), 517–527 (2005)
31. Matsumoto, T., Tani, J.: Goal-directed planning for habituated agents by active inference using a variational recurrent neural network. Entropy **22**(5), 564 (2020). DOI 10.3390/e22050564. URL `https://doi.org/10.3390/e22050564`
32. Mellmann, H., Schlotter, S.A.: Advances on simulation based selection of actions for a humanoid soccer-robot. In: Proceedings of the 12th Workshop on Humanoid Soccer Robots, 17th IEEE-RAS International Conference on Humanoid Robots (Humanoids), Madrid, Spain. (2017)
33. Mellmann, H., Schlotter, S.A., Blum, C.: Simulation based selection of actions for a humanoid soccer-robot. In: S. Behnke, R. Sheh, S. Sarıel, D.D. Lee (eds.) RoboCup 2016: Robot World Cup XX, pp. 193–205. Springer International Publishing, Cham (2016). URL `http://www.ais.uni-bonn.de/robocup.de/2016/papers/RoboCup_Symposium_2016_Mellmann.pdf`
34. Mellmann, H., Schlotter, S.A., Musiolek, L., Hafner, V.V.: Anticipation as a mechanism for complex behavior in artificial agents. In: J. Bongard, J. Lovato, L. Hebert-Dufrésne, R. Dasari, L. Soros (eds.) Proceedings of the ALIFE 2020: The 2020 Conference on Artificial Life, pp. 157–159. MIT Press Direct (2020). DOI 10.1162/isal_a_00314. URL `https://www.mitpressjournals.org/doi/abs/10.1162/isal_a_00314`
35. Mirza, N.A., Nehaniv, C.L., Dautenhahn, K., te Boekhorst, I.R.J.A.: Anticipating future experience using grounded sensorimotor informational relationships. In: ALIFE2008, pp. 412–419 (2008)
36. Nieuwenhuisen, M., Steffens, R., Behnke, S.: Local multiresolution path planning in soccer games based on projected intentions. RoboCup 2011: Robot Soccer World Cup XV pp. 495–506 (2012)
37. North, M.J., Collier, N.T., Ozik, J., Tatara, E.R., Macal, C.M., Bragen, M., Sydelko, P.: Complex adaptive systems modeling with repast simphony. Complex Adaptive Systems Modeling **1**(1), 3 (2013). DOI 10.1186/2194-3206-1-3. URL `https://doi.org/10.1186/2194-3206-1-3`
38. Olsson, L., Nehaniv, C., Polani, D.: From unknown sensors and actuators to actions grounded in sensorimotor perceptions. Connect. Sci. **18**, 121–144 (2006). DOI 10.1080/09540090600768542
39. Pezzulo, G., Butz, M.V., Castelfranchi, C., Falcone, R. (eds.): The Challenge of Anticipation: A Unifying Framework for the Analysis and Design of Artificial Cognitive Systems, *Lecture Notes in Artificial Intelligence (LNAI)*, vol. 5225. Springer Berlin Heidelberg, Berlin, Heidelberg (2008)
40. Pezzulo, G., Butz, M.V., Sigaud, O., Baldassarre, G. (eds.): Anticipatory Behavior in Adaptive Learning Systems: From Psychological Theories to Artificial Cognitive Systems, *Lecture Notes in Artificial Intelligence (LNAI)*, vol. 5499. Springer Berlin Heidelberg, Berlin, Heidelberg (2009)
41. Pico, A., Schillaci, G., Hafner, V.V., Lara, B.: How do i sound like? forward models for robot ego-noise prediction. In: 2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob), pp. 246–251 (2016). DOI 10.1109/DEVLRN.2016.7846826
42. Rasmussen, C., Dupont, Y.L., Madsen, H.B., Bogusch, P., Goulson, D., Herbertsson, L., Maia, K.P., Nielsen, A., Olesen, J.M., Potts, S.G., Roberts, S.P.M., Sydenham, M.A.K., Kryger, P.: Evaluating competition for forage plants between honey bees and wild bees in denmark. PLOS ONE **16**(1), 1–19 (2021). DOI 10.1371/journal.pone.0250056
43. Röfer, T.: Cabsl – c-based agent behavior specification language. In: H. Akiyama, O. Obst, C. Sammut, F. Tonidandel (eds.) RoboCup 2017: Robot World Cup XXI, pp. 135–142. Springer International Publishing, Cham (2018)

44. Rosen, R.: Anticipatory Systems in Retrospect and Prospect, pp. 537–557. Springer US, Boston, MA (1991 (original 1979)). DOI 10.1007/978-1-4899-0718-9_39. URL https://doi.org/10.1007/978-1-4899-0718-9_39
45. Rosen, R.: Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations. Springer New York, New York, NY (2012 (orignal 1985)). DOI 10.1007/978-1-4614-1269-4_6. URL https://doi.org/10.1007/978-1-4614-1269-4_6
46. Schillaci, G., Hafner, V.V., Lara, B.: Coupled inverse-forward models for action execution leading to tool-use in a humanoid robot. In: Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI '12, p. 231–232. Association for Computing Machinery, New York, NY, USA (2012). DOI 10.1145/2157689.2157770. URL https://doi.org/10.1145/2157689.2157770
47. Schillaci, G., Hafner, V.V., Lara, B.: Exploration behaviors, body representations, and simulation processes for the development of cognition in artificial agents. Frontiers in Robotics and AI **3**, 39 (2016)
48. Schmickl, T., Crailsheim, K.: Costs of environmental fluctuations and benefits of dynamic decentralized foraging decisions in honey bees. In: C. Anderson, T. Balch (eds.) The 2nd International Workshop on the Mathematics and algorithms of Social Insects Proceedings, pp. 145–152. Georgia Institute of Technology, Atlanta, GA (2003)
49. Seeley, T., Camazine, S., Sneyd, J.: Collective decision-making in honey bees: How colonies choose among nectar sources. Behavioral Ecology and Sociobiology **28**(4), 277–290 (1991). DOI 10.1007/BF00175101
50. Stefanec, M., Oberreiter, H., Becher, M.A., Haase, G., Schmickl, T.: Effects of sinusoidal vibrations on the motion response of honeybees. Frontiers in Physics **9** (2021). DOI 10.3389/fphy.2021.670555. URL https://www.frontiersin.org/article/10.3389/fphy.2021.670555
51. Taliaronak, V., Mellmann, H., Hafner, V.V.: Simulation of interactions between beehives. In: Proceedings of the 29th International Workshop on Concurrency, Specification and Programming (CS&P 2021), pp. 106–112 (2021). URL http://ceur-ws.org/Vol-2951/
52. Tisue, S., Wilensky, U.: Netlogo: A simple environment for modeling complexity. In: C. Anderson, T. Balch (eds.) The International Conference on Complex Systems. Boston, MA, USA (2004)
53. Watkins, C.J.C.H., Dayan, P.: Q-learning. Machine Learning **8**(3), 279–292 (1992). DOI 10.1007/BF00992698. URL https://doi.org/10.1007/BF00992698
54. Winfield, A.F.T., Hafner, V.V.: Anticipation in robotics. In: R. Poli (ed.) Handbook of Anticipation: Theoretical and Applied Aspects of the Use of Future in Decision Making, pp. 1–30. Springer International Publishing, Cham (2018). DOI 10.1007/978-3-319-31737-3_73-1. URL https://doi.org/10.1007/978-3-319-31737-3_73-1