

Master thesis on Cognitive Systems and Interactive Media

Universitat Pompeu Fabra

Laboratory of Synthetic Perceptive, Emotive and Cognitive Systems (SPECS)

Homeostatic and allostatic principles for behavioral regulation in desert reptiles:

A robotic evaluation

Student: Tue Ngo

Supervisor: Oscar Guerrero Rosado

Co-supervisor: Ismael Tito Freire González

July 2021



**Universitat
Pompeu Fabra**
Barcelona

Abstract

The survival of living organisms is accomplished through regulatory behaviors. In the case of the specialist Namibian sand-diving lizards, those behaviors are modulated not only in response to external events in their immediate ecological context but also to satiate their internal drives. Both homeostasis and allostasis are endogenous processes responsible for maintaining the internal stability of animal physiological variables, in which allostasis serves as a controller coordinating multiple homeostatic subsystems. By performing homeostatic and allostatic regulation, the desert-adapted lizard can avoid extreme temperatures while being able to acquire limited resources. Moreover, a living organism can make behavioral adaptations to cope with chronic stressful situations presented by its unstable environment. Yet how multiple internal states are processed and calibrated during those processes has not been fully clarified due to how inconsistent “allostasis” is understood and applied in numerous researches. We concentrate allostatic control to behavioral adaptation without anticipation to highlight that homeostasis and allostasis offer complementary procedures to withstand immediate instability. This study integrates homeostatic and allostatic regulatory mechanisms based on interoceptive and exteroceptive cues into a simulated mobile robot replicating the lizard’s sand-diving and foraging behaviors. To implement drive competition and conflict resolution in the animal’s brain, we propose a computational model based on the interaction between the brainstem’s medial Reticular Formation and the hypothalamus. Such a bio-inspired system is capable of action selection, and thus, can generate complex behaviors upon stimuli received from both the environment and the agent’s internal states. Finally, we evaluate the robot’s performance under capricious environmental settings. Our results support a dynamic, reconfigurable hierarchical organization of internal drives as an essential feature of sufficient regulation that ensures a healthier constitution.

Keywords: Behavioral Regulation, Allostasis, Homeostasis, Stress, Allostatic Load, Action Selection, Reticular Formation, Cognitive Architecture, Biomimetic Robotics.

Table of Contents

1. Introduction	4
1.1. Problem statement	4
1.2. State of the art	6
1.2.1. Homeostatic and allostatic principles	6
1.2.2. Allostatic load and stress response	11
1.2.3. Neural substrate for action selection	14
1.2.4. Behavioral repertory of desert reptiles	16
2. Methodology	18
2.1. Research question and hypotheses	18
2.2. Model implementation	19
2.2.1. World – Self – Action paradigm	19
2.2.2. Gradient and non-gradient-based subsystems	21
2.2.3. Interoceptive and exteroceptive cues for adaptation	24
2.2.4. Mechanism for action selection	26
2.3. Technical description	27
2.4. Experimental design and setup	29
3. Results	32
3.1. Static and dynamic hierarchy of motivational drives	32
3.1.1. Results for homeostatic control	32
3.1.2. Results for allostatic control	36
3.2. Interoception and exteroception-based adaptive agents	38
4. Discussion	42
5. List of figures	44
6. List of tables	45
7. Bibliography	46

1. Introduction

1.1. Problem statement

The natural behavior of animals reflects the adaptive coupling between the organism and its ecological niche. Their complex behavioral repertoires go beyond simple reflexes of the opportunities and threats that the environment provides, but they are also not pure results of learning and evolution. More precisely, a living organism's physical body has the wisdom to constantly perform behavioral adjustments to handle internal disturbances that external environmental perturbations would otherwise cause.

According to Hull (1943), drive reduction is the primary force behind behaviors [1]. Drives are endogenous and instinctual urges that arise due to physiological or psychological needs of biological beings. Behavioral regulation in animals is one of many ways to reduce those impulses. Various recent studies in regulatory behaviors [2, 3] proposed homeostasis as the primary assessment to classify motivational drives. Homeostasis refers to the “steady state” of internal physiological systems that are essential for organismic viability [4]. Evidence of homeostasis can be observed by analyzing the diurnal behaviors of the sand-diving lizards inhabiting the Namib desert in Southern Africa [5, 6]. Under extreme temperatures in the desert, the lizards engage in complex self-regulatory behaviors such as diving beneath or rising above the sand surface to maintain a desired body heat of 30°C. Besides temperature, numerous homeostatic systems operate simultaneously to control other needs (e.g., water, food, sex, or anti-predation). At a meta level, evolved biological systems are driven to move between homeostatic systems and maintain a dynamic equilibrium of multiple needs being fulfilled. Allostasis serves this purpose by counteracting fluctuations through changes [7].

Sterling and Eyer (1988) originally coined the term “allostasis” and elaborated its beneficial roles in [7]. One is that it permits an adequate matching of resources to demands. This is done through a continuous re-evaluation of needs and a regular adjustment of permissible homeostatic thresholds in order to generate motivational tendencies that are not too high or too low for whatever is happening outside the body [7]. The second claimed advantage of allostasis is its capacity to anticipate upcoming needs and make necessary responses in advance. Under this allostatic regime, the body benefits from the animal's capacity to learn from experience [7]. Several succeeding studies [2, 3, 8, 9, 10, 11] considered an allostatic controller that supervises and orchestrates multiple homeostatic subsystems. However, they interpreted the characteristic features of allostatic control inconsistently. In [8, 9], allostasis adjusts the favorable values (i.e., the setpoints) of individual homeostatic subsystems, plus integrates and weights their outputs to better regulate multi-perspective behaviors. Lee et al. [10] likewise referred to allostasis as changing setpoints for homeostatic recovery. Those

alterations are made in response to chronic stressful events or poorly regulated responses without any clue to inform the animal that an external fluctuation is about to come. On the other hand, allostasis is proposed by [11] to be the prospective preventions in which the brain consolidates the stored “prior knowledge” and incoming sensory inputs to anticipate what resources will most likely be needed, and thus, minimizes upcoming discrepancies between the currently sensed and desired values of homeostatic subsystems. Nevertheless, those compensatory adjustments require frequent updates on the intensity of each drive (i.e., the difference between the actual and desired values of the corresponding homeostatic system). Vouloutsi et al. [2] additionally suggested that the relative importance (i.e., the priority levels) of motivational drives play an equally important role as their intensities in allostatic control. Moreover, the priority levels of drives could be ordered hierarchically in a way that moderately matches Maslow’s pyramid [12], where higher needs would only be considered if those constituting the base are fulfilled.

Due to how ambiguous the term “allostasis” is understood and applied in those various researches, the calibration and coordination of multiple motivational drives remain unclear. Therefore, it appears necessary for us to review the state of the art in allostatic control and articulate a more concrete definition of allostasis. Our proposal for the terminology would be: whereas homeostasis indicates the local, primitive feedback loops involving autonomic reflexes that resist physiological disturbances, allostasis ameliorates internal perturbations through an integrative, hierarchical orchestration of multiple homeostatic subsystems being dynamically reorganized. In this study, we concentrate the scope of allostasis to adaptation without any predictive regulation to highlight that homeostasis and allostasis indeed offer complementary procedures for the mutual goal of coping with immediate instability.

This study follows a synthetic methodology of convergent validation, starting with fine-tune a model of homeostatic and allostatic control proposed by [13]. Our main objective is to enhance the efficiency of behavioral regulation by empowering the control system with a reconfigurable hierarchy of motivational drives. The result will be subsequently integrated into a mobile robot capturing essential properties of the observed behaviors as navigation profiles. To build an agent which robustly operates in a complex environment, an efficient strategy is to reverse-engineer the existing biological structures that have already solved the problem. The hypothalamus is known to be the convergent center of interoceptive signals, regulating many effectors of homeostasis [14]. We look for a neural substrate for behavior selection that interacts with both ascending hypothalamic and descending reticulospinal systems, and the brainstem’s medial Reticular Formation seems to be the best candidate [15, 16]. The medial Reticular Formation and its interaction with the hypothalamus serve as the starting point for designing our proposed mechanism of action selection [17, 18].

Our proof of concept will be examined entirely through robotic simulations. We expect the mobile robot to replicate the sand diving and foraging behaviors within an environment simulating the desert lizard's ecosystem. Thus, we can prove the hypotheses regarding the efficiency of our approach on how animal behaviors are regulated. If the simulated robot performs poorly, then our model is likely to be incorrect.

1.2. State of the art

1.2.1. Homeostatic and allostatic principles

In his drive reduction theory, Hull (1943) defined the term “drive” as an unbalanced state that governs and affects behaviors. Primary drives are often innate and pertain to the basic needs of animals, such as warmth, hydration, nutrition, self-preservation, and reproduction. Under aversive (e.g., excessive temperature) and pleasurable (e.g., food or sex) conditions, the organism creates tensions or arousals that need to be reduced. Animals are motivated to carry out regulatory behaviors that lead to the reduction of those drives [1].

Hull developed his theory around the concept of homeostasis which refers to the optimal processes used to reach and maintain internal states at fairly constant and stable levels [4]. The steadiness of internal states is reflected through the constancy of “le milieu intérieur”, i.e., the extracellular fluid environment that envelops the cell [5]. The goal of homeostatic control is to maintain the balance and consistency of this interior milieu by strictly ensuring essential physical and chemical conditions (e.g., body temperature, pH, oxygen tension, or glucose level) being kept within viable homeostatic ranges. In this balanced state, certain motivational drives encounter appropriate stimuli that satisfy their needs. Both the absence and excessive presence of a stimulus can lead to an unbalanced state of the corresponding internal system, and thus, create a tendency to take a countervailing action. This capacity to cancel fluctuations out with regulatory responses helps the living organism adapt to its unique habitat. An enlightening example of the Namib desert lizard can be found in [6]. In general, reptiles lack an internal mechanism of metabolic heat production as endotherms, which results in their reliance on external heat [5]. Hence, they perform thermoregulation exclusively through interactions with the environment. When the surface temperature is at 30°C, the lizard emerges from the sand and presses its ventral surface against the substrate, having its body heat restored. Once the surface temperature starts to increase, it performs a “thermal dance” – lifting its feet off the hot sand – in order to prevent the limbs from burning. Should it get extremely hot (i.e., 40–50°C), the animal digs into the sand dune where temperatures are cooler [6]. In terms of homeostasis, the ectothermic lizard needs to sustain a desired temperature value of 30°C. Lower or higher temperatures motivate the animal to behave in such a way that brings the actual body heat back to 30°C. However,

homeostasis cannot be the only regulatory mechanism since the premise of it does not precisely address the primary goals of animal behavioral regulation. Regulatory behaviors are meant to promote survival and reproduction. Sustaining the constancy of any of the distinctive homeostatic systems cannot be the ultimate concern of the sand-diving lizard. Under the harsh living conditions in the desert, the animal needs to develop an optimal strategy to not only escape from thermal stress but also survive far from obvious resources [5, 6]. This is where the notion of allostasis came into use. Instead of preserving constancy inside closed homeostatic loops, allostatic control achieves global stability by changing the dynamics and priorities of different homeostatic subsystems pertinently to environmental changes [7, 8]. As a result, multiple motivational drives are satiated, taking into account the consequences of actions. In the case of the sand-diving lizards, being beneath the sand precludes the acquisition of food and water. The organism is able to determine whether it should prioritize hydration over preserving its body heat, or how much danger should it tolerate while searching for resources. When the lizard is replete with water content, it will spend less time on the surface of the sand, prioritizing self-protection. On the contrary, if the desert lizard is profoundly thirsty, it will take a higher than normal risk of predation by venturing onto the surface, avoiding an eventual death [6].

With these understandings of biological concepts, we consult several studies that have been conducted following the same methodology of a biomimetic agent-based architecture built upon computational processes. An inspiring system of how a motivational drive interacts with and is expressed through behavior is shown in [19]. Their framework consists of a motivation system (including the agent's internal drives), a behavior system (containing a set of regulatory behaviors), a perceptual system (extracting salient external stimuli), and a motor system (implementing motor skills). Each motivational drive is modeled as a specific transducer process that is active when its activation level exceeds a threshold and expresses itself through a regulatory behavior. All drives and behaviors are organized in such a way that the agent tries to achieve homeostatic balance and preferentially passes activation to some behaviors over others. In this respect, [2] developed a similar yet more progressive architecture. Their chosen drive set was different from ours since they explored behavioral modulation in the context of human-robot social interaction. Still, the most important improvement of [2] compared to [19] is that a certain drive is not mapped to a specific behavior on a one-to-one basis. Rather, each drive is managed by a homeostatic controller that defines its ideal range and evaluates whether its actual value is within that range. In addition, the authors proposed in [3] that the homeostatic controller then classifies the motivational drive into one of the following states: "under homeostasis", "homeostasis", and "over homeostasis". Both "under homeostasis" and "over homeostasis" states trigger

behavioral regulation. Figure 1 depicts an unspecified homeostatic system with maximum and minimum bounds. The system is classified as balanced while the actual value (aV) at a given time stays within the predefined desired value (dV) range. The agent only exhibits the modulation when the aV goes below the minimum limit or above the maximum limit.

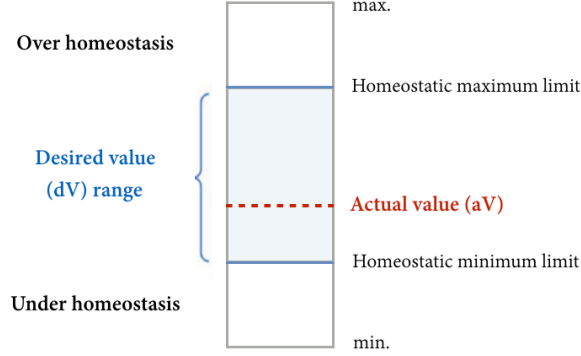


Figure 1: **Homeostatic control.** Keeps the actual value (aV) of the controlled parameter within the desired value (dV) range. The dV range is determined by the homeostatic limits. Reproduced from [2, 3, 13].

On top of the homeostatic assessment of each drive, there is an allostatic controller which constantly monitors the intensity of motivational tendencies and their levels of priority [2]. The ultimate goal of allostasis is to inform the animal about which need to fulfill foremost. Therefore, a more complicated yet more effective regulation system of flexible trade-offs is empowered. We modify the control mechanism described in [2, 3] into a new one suitable for our research. Key components are adapted to animal's biological behaviors, following suggestions from [13]. Figure 2 demonstrates how an organism evaluates its internal states based on the principles of intensity and priority. Four typical homeostatic subsystems (anti-predation, thermoregulation, thirst, and hunger) are arranged from left to right according to their priority levels. The dV ranges of different homeostatic subsystems also differ. The intensity of the tendency to bring aV back to balance is calculated by the distance from the aV (being outside the dV range) to the closest homeostatic boundary. Light and dark blue dashed lines in the thirst homeostatic subsystem represent different aVs , with the dark blue line (2) indicating more intense demand than the light blue one (1). The “hide” behavior is meant to satisfy the anti-predation drive, and the “explore” behavior fulfills the thirst drive. When anti-predation and thirst drives contradict each other, the intensity of anti-predation need will be compared to the intensity of thirst. In the first case demonstrated by the light blue dashed line (1), both tendencies are equally urgent. As a result, the action selected will be based on the priority principle which favors safety from predators over hydration. The animal will hide beneath the sand and silence its slight need for water. However, if the thirst drive with lower priority has a higher intensity, the living organism will prioritize the satiation of thirst since its survival is fatally endangered. That being so, in the second case

demonstrated by the dark blue line (2), the tendency for water acquisition will suppress the need for self-protection, even in the presence of predators.

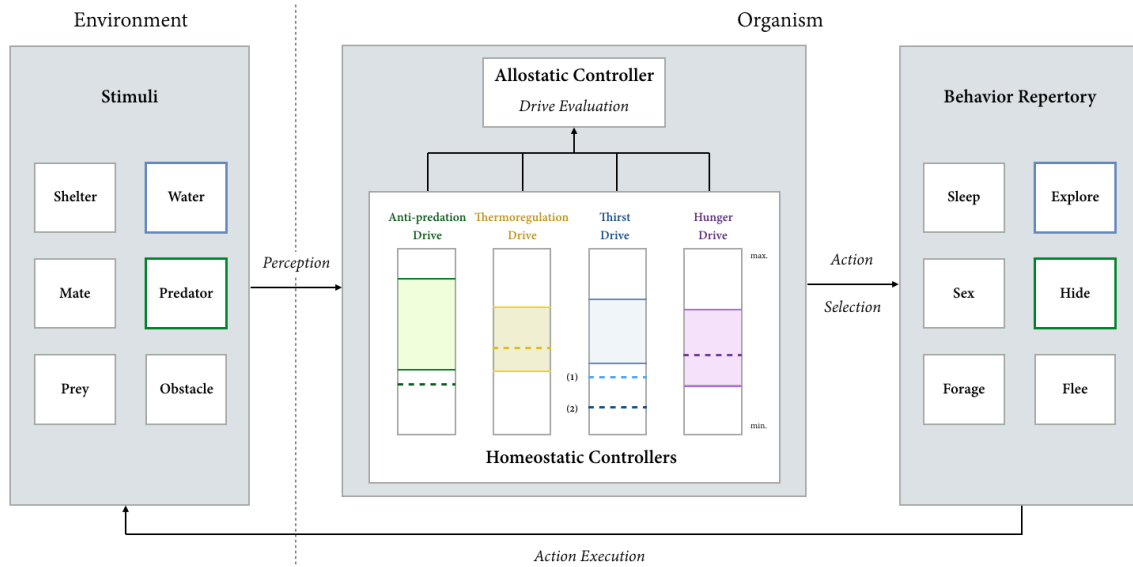


Figure 2: **Detailed diagram of behavior selection.** Internal drives are evaluated by means of intensity and priority. Case (1) prioritizes anti-predation over thirst, while case (2) prioritizes thirst over anti-predation. Reproduced from [2, 3, 13].

While the way we estimate the intensity of a certain internal drive seems to be apparent, how the allostatic controller determines the level of priority remains unclear. We support the idea from [2] that the priority levels of different motivational drives will likely form a pyramid approximating Maslow's hierarchy of needs [12]. The most enduring concept of Maslow is probably the hierarchical arrangement in which some drives take precedence over others. Maslow's proposal that the fundamental needs are independent and multiple is aligned with the existence of numerous homeostatic subsystems – determined by different dynamics, stimuli, and neural substrates. Even though the basic foundations of the classic pyramid are worth preserving, it was specifically built to classify the drives and motives of humans, not those of animals. Maslow himself extended the homeostatic principle only to all preceding four levels due to their deficit nature [12]. On the contrary, self-actualization needs in humans do not involve homeostasis since they are likely to be nourished once engaged. Hence, this layer should not be considered in studies regarding homeostasis and animal behavior. A revised hierarchy for animal's needs was introduced in [20], integrating theoretical and empirical evidence at the interface of evolutionary biology. We incorporate this renovated hierarchy (Figure 3) in our model, plus propose two main modifications. First, in Maslow's classic pyramid, all physiological drives are depicted as equal in terms of preference, while in reality, they are not. For example, excretion is more important than water consumption, and thermoregulatory needs cannot oppose the demand for oxygen.

Furthermore, among immediate metabolic drives, maintaining body temperature outwits the urges for water and food [21]. Hence, we propose that each layer depicted in Figure 3 should contain needs with different priority values. Second, contemporary theorists pointed out that fundamental drives are not intrinsic but goal-oriented, with their dynamics being calibrated in response to situational threats and opportunities [20]. Thus, the layers of the hierarchy used in our study should be reconfigured corresponding to immediate external triggers. This last point evokes the idea of an adaptable hierarchical organization of needs that could serve as a valuable extension to the previous works in allostatic control.

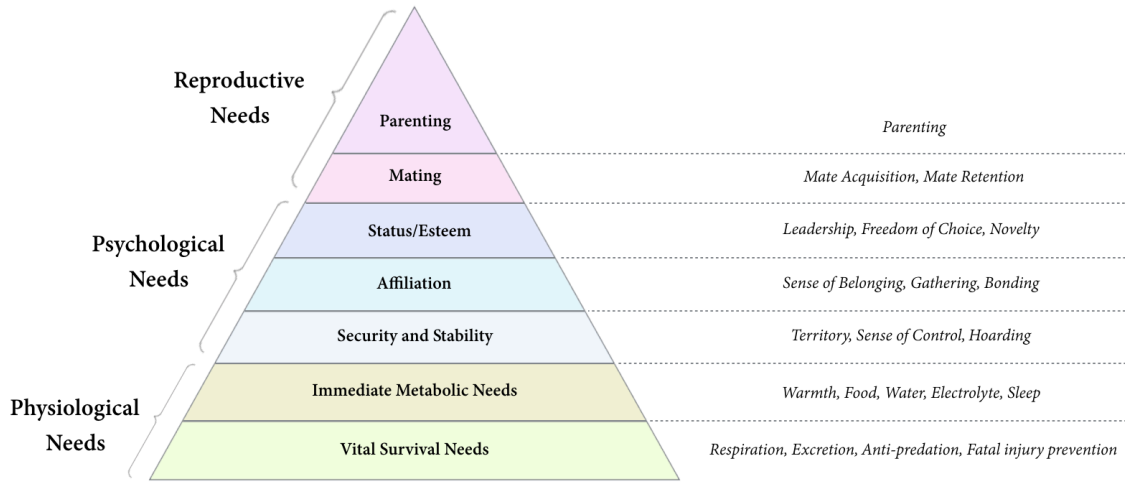


Figure 3: **The animal hierarchy of needs.** The activation of a drive-based goal will be triggered whenever relevant environmental cues are salient. Reproduced from [20].

Despite various successful models exploiting concepts of allostasis, there are only a few benchmarks for such biomimetic control systems. The most notable one might be the work by Sanchez-Fibla et al. [8, 9]. They implemented homeostatic and allostatic principles in a mobile robot imitating rat behaviors, making [8, 9] exceptionally similar to our approach. Behaviors of real rodents chosen as their project’s benchmarks were recorded through an open field test and include free exploration, homing, reward localization [8], plus advanced capacities such as path finding [9]. The authors assessed a simulated robot and a physical one steering in different experimental arenas. The model was validated by comparing the agents’ outputs with real rat behaviors. Drive-based self-regulation was decomposed into a minimal set of homeostatic subsystems approximating the observed behaviors. Those subsystems are depicted as gradients, and the desired values portray favorable locations in such landscapes. For instance, the “arousal” gradient has a maximum value in the center that drops to zero towards the experimental boundaries, indirectly measuring arousal levels as distances. The “security” gradient fixes the desired value at a preferred corner, and the “cue” gradient changes itself dynamically. Regarding allostasis, Sanchez-Fibla et al. argued

that setpoints of the homeostatic gradients can be adjusted through a probability switching mechanism to achieve stability at a behavioral meta-level [8]. Besides adjusting the desired values, allostatic control is responsible for bringing the agent closer to the desired location through the integration of the gradient-dependent information and its mappings onto the motor system. The motor contribution of a given gradient to navigation is influenced by the difference between the actual (aV) and desired (dV) values and a fixed weighting factor k that is different for every gradient. Such implementation is relevant to how we previously established the principles of intensity and priority. Sanchez-Fibla et al. also discussed how allostatic control could be applied to perform complex path planning tasks. The resulting regulation was reflected not only in the internal stability that the robot achieved but also in its ability to form an optimal path toward a target [9].

1.2.2. Allostatic load and stress response

Although Sanchez-Fibla's work is chosen as the valuable benchmark for our study, we see some areas where the implemented model can be extended. Replicating the model in [8, 9] together with overcoming its limitations will certainly support the development of a more comprehensive and beneficial architecture. A potential extension of Sanchez-Fibla's model could be the generalization of allostatic control from simple corrections of already-sensed mismatches between the desired and actual homeostatic values to adaptations that inhibit those erroneous outcomes from growing dangerously large, especially within a stochastic environment with limited observability of resources.

It is worth clarifying whether allostasis necessarily implies forecasting future needs through learning and preparing to fulfill needs before they arise, and to what extent can allostasis provide adaptations. From a control-theoretic perspective, the homeostatic objective is to minimize the discrepancy between favorable and actual interoceptive sensations [22]. This error signal is usually referred to as “free energy” in studies [22, 23] adopting the notion of interoceptive inference. In interoceptive inference processes (including homeostasis and allostasis), the living systems encode representations of the body's physiological states and provide vital information about how well the living body maintains its inner environment's steadiness. The main goals of interoceptive inference include minimizing error signals in terms of both magnitude and frequency, and then utilizing those errors in reducing future ones [22]. The review in [23] categorized current approaches on minimizing “free energy” into three broad classes: behavioral, teleological, and diachronic allostasis. While the two latter ones shift the focus to a broader perspective, emphasizing beforehand predictions of future errors and higher-level implementation of flexible action policies, the “behavioral” approach strikingly retains allostasis as a reactive mechanism equivalent to the automatic corrective responses orchestrated by homeostatic systems. On this account, homeostasis is

a group of first-order negative feedback loops involving autonomic regulation constituted by the dynamic interplay of each homeostatic module's inputs and outputs. If homeostatic control fails to stabilize essential physiological variables within viable boundaries, allostatic control intervenes to reconfigure the homeostatic systems' dynamics and rankings without any anticipatory behavior in the service of any particular homeostatic module [23]. Thus, although this second-order arrangement lacks the capacity to predict and offset deviations before they occur, it still can offer adaptations to the long-term consequences caused by unresolved internal perturbations. Observations of such responses found in nature include: seeking shelter before night falls, hoarding before winter, longer sleep cycles, or elevated gluconeogenesis [24]. Indeed, allostasis is proposed by many studies [24, 25] to be the underlying process in which animals carry out the information integration and processing of physiological conditions, preceding experiences, and habitat configurations in order to avoid damaging situations. Unpleasant experiences – either derived from predictable (e.g., seasonal shortage of resources) or unpredictable changes (e.g., natural disasters) in the animal's habitat – are perceived as threats to survival, cause stresses, and significantly force behavioral changes. Thus, living organisms can adapt to capricious environments [25].

The aforementioned observations embraced allostasis as a promising concept for studying complex relations between behavior, stress, and allostatic load in coping strategies with the hazards presented by uncertain environmental conditions [24]. Stresses are stimuli inducing changes in homeostasis (i.e., allostasis) of internal organizations constituting the autonomic nervous system (e.g., blood pressure), metabolic hormones (e.g., cortisol, insulin), and pro/anti-inflammatory cytokines [10]. The cumulation of slight yet repetitive stresses can lead to long-standing consequences as they constantly and excessively inhibit the homeostatic recovery processes. When increased need generates a homeostatic “error”, the first-order feedback mechanism works on minimizing that error. However, the desirable stimuli might not be available by then, and the time needed for correction might be exhaustively long. Allostatic load is thus the accumulating physiological burden of adaptation to prolonged stress. Behavioral responses modulated by allostatic load are initiated by a stress factor, sustained for an appropriate time interval, and then shut off when the stress is terminated [10]. Several conditions might lead to allostatic loads, such as gradually elevated stress, the lack of adaptation to frequent stress factors, elongated responses after a threat is past, and inadequate responses that trigger unnecessary compensations in other behaviors [26].

However, how allostatic load is computed is still poorly elaborated across the literature. We start with the premise that the most appropriate technique to minimize homeostatic errors could be to modify interoceptive parameters so that they correspond with the perceived stress level of environmental and physiological conditions [22]. According to Lee et al. [10],

if the frustrating situation persists, the living system might require the relevant homeostatic systems to update their setpoints (i.e., desired values) for future adaptation, and allostatic load is defined as the difference between the newly defined and previous setpoints. Such a dynamic adjustment of setpoints was implemented by Sanchez-Fibla et al. in [8, 9] through a probability switching mechanism, and it approximately explains adaptation in the case of several homeostatic modules interacting with each other.

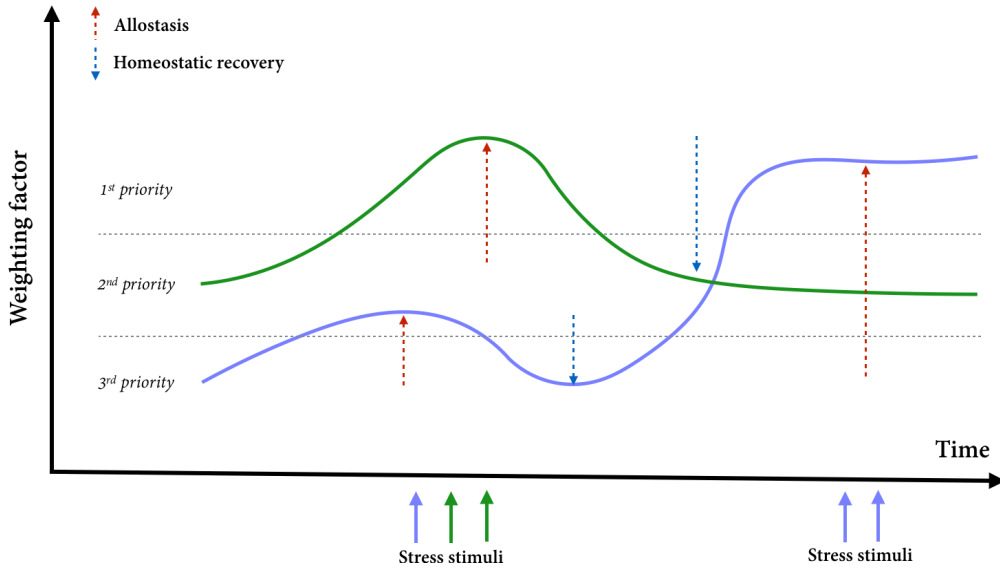


Figure 4: **Dynamic configuration of the weighting factor.** Blue and green solid lines represents related stress stimuli. Chronic stress increases the relative priority and motivates the organism towards fulfilling the need (red dashed lines). When the need is satisfied, the homeostatic subsystem is relaxed (blue dashed lines) and the relative priority decreases. Reproduced from [10, 11, 26].

Even so, an absolute reconfiguration from an initial desired value to an “allostatic” one leaves us in doubt. We look for an alternative methodology to compute allostatic load that is consistent with its cumulative nature [26] and a cyclic regulatory fashion of “arousal” and “relaxation” as described in [11]. We additionally propose that allostatic load is not only the label for the discrepancy between the initial and “allostatic” values but also the mechanism which results in that adjustment. Allostatic load should be calculated in such a way that requires the continuous storing of the growing stress for each stimulus (e.g. temperature, food, or water). In this respect, whenever homeostatic subsystems fail to resolve internal perturbations caused by stress, the allostatic load value increases. The organism is driven towards fulfilling the need, and when the homeostatic subsystem recovers, the allostatic load value decreases [11]. Furthermore, not the desired value but the priority level of the homeostatic subsystem is modulated by means of allostatic load. Allostasis is often applied to explain a wide variety of behavioral adaptations, from fluctuations in blood pressure to changes in food consumption. When the allostatic adjustment is going to be maintained for

a long period, such as the blood pressure case, modifying homeostatic setpoints apparently makes sense. In the case of proximate stimuli such as food or water, the reconfiguration of priority levels seems to be a reliable motivation-based alternative. Figure 4 gives an example of two arbitrary homeostatic systems with their priority levels being rearranged in response to stress stimuli and increased allostatic load. This direction satisfies the prerequisite that allostatic control can sense an average value for arterial pressure and gradually moves it upward. At the same time, it goes aligned with the notion of a reconfigurable hierarchy of motivational drives elicited in section 1.2.1.

1.2.3. Neural substrate for action selection

In addition to the missing adaptation part, the work in [8, 9] has another limitation related to the ambiguous borderline between drive assessment and behavior selection. Allostasis should be distinguished with the resolution of conflicting needs. In reality, action selection does not necessarily imply adjusting the internal systems' parameters or minimizing future mismatched outcomes. There is an immediate trade-off between water acquisition and self-preservation when an animal is thirsty and hunted by a predator at the same time, and the decision of which action to trigger should also consider the salience of the water stimuli or the disposition of predators. Therefore, action selection should be treated as a consequent yet distinct process from allostasis. Referring to the previous literature review, we define four requirements for an effective central selection system. First, the system must integrate information from both internal and external states. Second, it is capable of interpreting the dominant motivational drive from that information. Third, it has an internal configuration for behavior representation and conflict resolution. Fourth, it generates outputs that allow selected actions to express themselves. Those requirements can be partially satisfied by the medial Reticular Formation (mRF) located in the brainstem and its interaction with other neural structures, especially the hypothalamus.

As stated by Plaff (2016), there should be a primitive, elementary, and universal neuronal “force” underlying all motivated behaviors. This powerful “force” – called central nervous system (CNS) arousal – is claimed to be one of the essential capacities of the brain as a central action selection device [15]. Generalized CNS arousal switches on ascending and descending systems in response to stimuli, and in consequence, facilitates the initialization of any behavior. The relationship between arousal, motivational drives, and consequential behaviors could be expressed mathematically as a linear equation [15]. In short, arousal is a vector formulated upon the internal and external states of the system, and the magnitude of this vector is capable of modulating behaviors. Humphries et al. [16] proposed that the best neural substrate to be assigned with the task of generalized CNS arousal control – and thus, action selection – is the mRF. The mRF is a group of interconnected nuclei located in

different parts of the brainstem. Regarding external connections to other neural structures, the mRF includes ascending neuromodulatory pathways (i.e., ascending reticular activating system responsible for maintaining behavioral arousal and setting the overall motivational state of the animal) to the cerebral cortex and descending reticulospinal tracts (in charge of somatic motor control) to the spinal cord [15, 16]. The reticular activating system projects anteriorly to the hypothalamus and receives input information from every sensory organ available to the organism [14]. The hypothalamus receives the neuroanatomical projections from the reticular activating system and modulates physiological and behavioral effectors of homeostasis. Behavioral regulation can be achieved through a hierarchical arrangement in which hypothalamic sensing could be linked to the mRF [14].

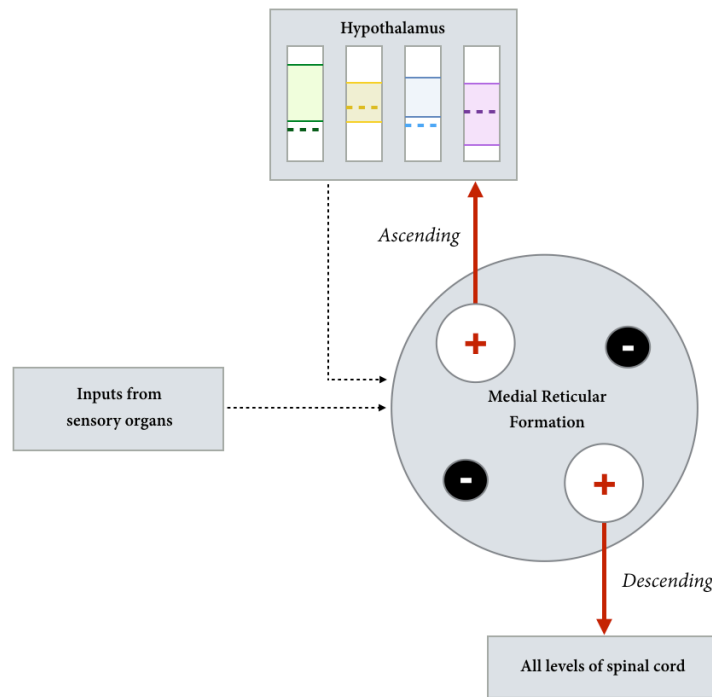


Figure 5: **Computational model of the medial Reticular Formation.** The mRF incorporates inputs from both sensory organs and internal homeostatic systems (black dashed arrows). Projection neurons (white large circles) excite arousal pathways (red arrows). Interneurons (black small circles) inhibit projection neurons. Reproduced from [13, 16, 17, 18].

In terms of internal circuitry, the authors of [16, 17] described the predominant neuron in the mRF – referred to as the projection neurons – as having a giant body and symmetric dendrites. A computational model is reproduced from [16, 17, 18], taking inspiration from the integration between the mRF and the hypothalamus (Figure 5). In this model, the projection neurons are arranged in a series of stacked clusters, together with another type of nerve cell called interneurons. Each cluster contains a combination of both types. Giant projection neurons make excitatory synaptic contacts with both ascending and descending

arousal pathways, while smaller interneurons inhibit the activation of projection neurons. Interneurons project exclusively within the cluster while projection neurons contact solely the neurons that belong to other systems (i.e., the hypothalamus and all levels of the spinal cord) [16, 18]. Projection neurons collect an array of unprocessed sensory inputs from all the ascending tracts of the brainstem and relay these afferents to the cortex. This endows the mRF the role of a brainstem “integrative core” capable of intrinsic (putting together homeostatic information) and extrinsic (reacting to sensory input) action generation. The mRF is also promoted as a “mode selector” which decides the global behavioral state of a vertebrate animal and influences regulatory actions [17]. The configuration of the mRF is thought of as a bottleneck necessary for converting massively concurrent and distributed information into constrained modes for action selection [17, 18].

1.2.4. Behavioral repertory of desert reptiles

As discussed above in section 1.2.3, arousal level could be calculated upon both external context and internal states of the robotic agent in the form of a vector with a magnitude capable of regulating behaviors. The resulting behaviors can vary in response to the same motivation, corresponding to different lengths of that arousal vector. Consequently, the animal can exhibit a rich and diverse behavioral repertory.

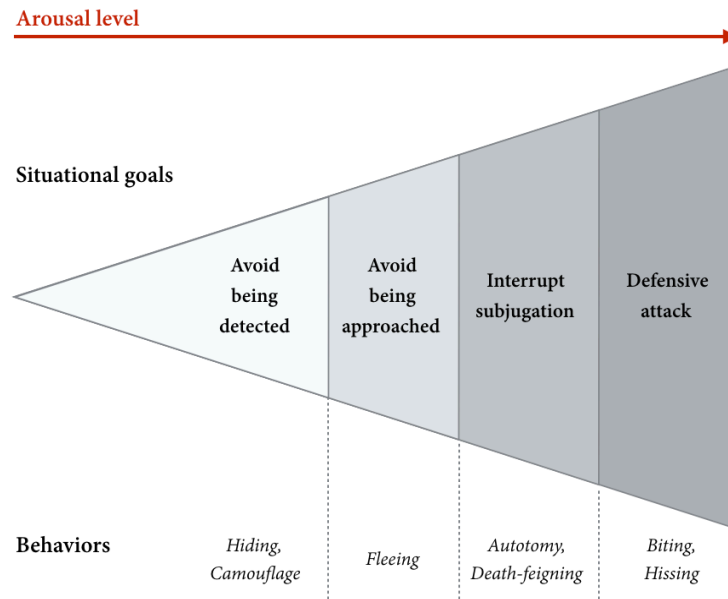


Figure 6: **Behavioral repertory for motivational tendency.** Regulatory behavior depends on arousal level. Reproduced from [13, 28].

Reptiles perform behavioral changes based on the cost of interrupting current activities (e.g., inaccessibility to resources) and the risk of stressful situations (e.g., the disposition of the predator or distance from a refuge) [27]. Most of the foraging and defensive tactics

exhibited by reptiles refer to possible resources and predators as visuospatial cues. Figure 6 arbitrarily presents a diversity of defensive behaviors with respect to the four main stages of a predatory event: detection/identification, approach, subjugation, and consumption [28]. During the initial stage of being detected and identified, lizards may try to stay safely out of reach from their potential predators. While hiding is the most obvious way to avoid detection, lizards additionally exploit confounding camouflage methods. If being identified as edible prey, similarly to other animal species, lizards perform locomotor escape (i.e., fleeing). After being approached, reptiles usually impede subjugation by discarding their own appendages to distract the predator or to elude its grasp. Tail autotomy is an extremely peculiar and common self-defense tactic in lizards. Additionally, the animals exhibit death-feigning in which they “play dead” by maintaining a frozen posture. Most notably, reptiles can aggressively and abruptly attack the predator (e.g., hissing, biting, or scratching) when they are severely threatened. Currently, there is no evident observation regarding reptilian tactics to avoid consumption [28].

The situational changes in behaviors can be explained by Merker’s selection triangle [29]. According to Merker, three coherent steps should be taken to adopt the most appropriate action. First, the living organism follows an evaluation process of internal states. Second, the target needs to be delicately weighted and selected. Third, the behavior congruent best with both the dominant internal states and the detected stimuli is selected. In this manner, a wide range of regulatory behaviors can be expressed depending on the arousal needed to trigger it. By computing generalized arousal as a vector, behaviors of the same category can also vary at different thresholds of the vector’s magnitude.

2. Method

2.1. Research question and hypotheses

In section 1.2.2, we raise the need to reach a comprehensive concept of allostatic control which integrates, reorganizes, and weighs multiple homeostatic subsystems. Critically, we propose that homeostatic regulation is supervised by allostatic control as a meta layer that persistently enables the organism to attain a fine coupling between internal needs and what the stochastic environment can offer [7]. Among the vast data that can be acquired during the perceptual processes, the preceding sensory events and the configuration of external surroundings seem to be the most informative ones [24]. According to [25], an unpleasant sensation encodes an impending threat to survival, while an appetitive one resembles an opportunity. We leverage the probability of environmental stimuli as the representation of threats and opportunities provided by the animal’s ecosystem within a larger time scale (in comparison to the time step in which the allostatic controller monitors and integrates homeostatic signals). The abundance or scarcity of resources as well as the frequency of alarming situations (e.g., excessive temperature, predatory events) can be examples of this variable. A model promoting adaptations based on the probability of environmental stimuli would be our main contribution to the previous works in allostatic control.

Regarding Sanchez-Fibla’s model [8, 9], besides the adjustment of the homeostatic desired values, in order to formulate the motor contribution of a specific homeostatic system, an invariant weighting factor k is applied uniformly during the agent’s lifetime. The value of this k factor can be seen as reflecting the position of the corresponding motivational drive within the hierarchical ranking organization explained in section 1.2.1 (Figure 3). The higher the importance of a motivational drive, the higher the value of the k factor, the more the drive is weighted, which results in a higher probability of the related action to be triggered. This implies a potential association between the hierarchy of internal drives and our introduced notion of environmental threats and opportunities.

Thus, we pose the research question regarding the relationship between the probability of environmental stimuli and the priority level of motivational tendencies. We envision that the sparseness of a favorable type of stimuli will weigh the associated drive to a greater extent than a setting in which the probability of those stimuli is high.

Moreover, we would like to investigate the relevance of environmental settings during drive evaluation and behavior selection processes, and how such information is used by living organisms to perform allostatic adaptations. As discussed, the relative priority level of a motivational drive requires a calibrating parameter representing cumulative perceived stress. This factor could be formulated in either an interoceptive or exteroceptive manner. The

exteroceptive cues for adaptation could straightforwardly be the chance of encountering a wanted stimulus, which correlates with the number of resources or events that are detected and stored during an appropriate time interval. On the other hand, allostatic load could be reliably integrated as an interoceptive calibrating factor, considering it as the disturbing cost of chronic exposure to unresolved threats. In this case, the value of allostatic load would be computed as the mean of accumulated drive intensities during a proper time span.

Thus, our three concrete hypotheses are:

H1. A reconfigurable hierarchy of drives with dynamic weighting factors allows the agent to achieve higher internal stability than a static one with constant weighting factors.

H2. An interoceptive agent configuring its hierarchy of drives based on allostatic load can regulate its behaviors as sufficiently as one based on external sensing of stimuli availability.

H3. The adaptation of the interoceptive agent is reflected in its behavioral self-regulation and correlates with the changing probability of environmental stimuli.

2.2. Model implementation

In summary, allostatic control in our model should be implemented to meet requirements regarding its capacity of integrating, reorganizing, and weighting motivational tendencies. It should permit an ongoing assessment of the matching level between internal states and external conditions as well as an adjustment of hierarchical weighting factors in response to changing probability of environmental opportunities and threats. The model built in [8, 9] has already offered some solutions, viz., gradient-based homeostatic loops, computed local perception, and the distribution of motor energy to behaviors based on the intensities of homeostatic subsystems. We plan to incorporate some of those solutions in our extension and include the physiological and neuroscientific evidence aforementioned in section 1.2.

2.2.1. World – Self – Action paradigm

We aim to decompose our model into sequential and hierarchical modules representing the fundamental functions of behavioral regulation (i.e., encoding, evaluating, and selecting). Such a layered structure is influenced by the Distributed Adaptive Control (DAC) cognitive architecture originally built by Verschure et al. [30] and adapted in many pieces of research in allostatic control [2, 31]. As per its description, DAC includes four tightly coupled layers (viz., somatic, reactive, adaptive, and contextual) stacked on top of each other according to their levels of complexity. The somatic layer comprising the sensors and actuators of the body, plus all physiological variables essential for its survival. The reactive layer serves drive reduction and contains predefined and stereotyped sensorimotor loops supporting basic functionalities (e.g., seeking, fleeing, etc.) and generating internal signals able to modulate

the subsequent layers. The adaptive layer allows the agent to deal with environmental uncertainty through learning from sensation-reflex couplets and shaping actions by both internal and external states. The contextual layer includes memory-based goal planning and policies that reinforce predictions. These layers can be grouped into three columns that represent the agent’s perception of the exterior world (i.e., “world”), its internal states (i.e., “self”), and actions that serve as the interface between those two (i.e., “action”). We center our work on DAC’s reactive and adaptive layers.

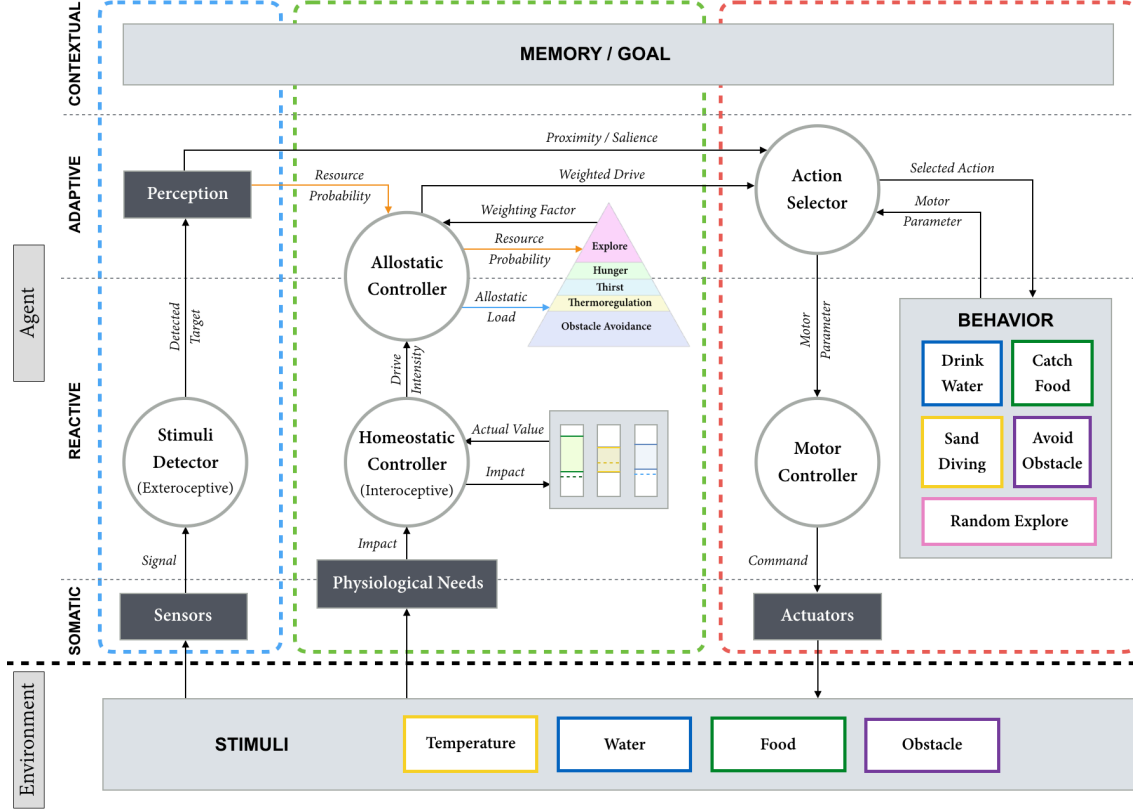


Figure 7: **DAC’s World – Self – Action framework.** The “world” column (left, blue) provides information from the environment. The “self” column (middle, green) senses, integrates, evaluates, and weighs internal variables. The “action” column (right, red) chooses the proper action based on the most salient stimulus and the most dominant drive, then generates motor commands to execute that selected action. Orange arrows are exclusively exhibited in exteroception-based adaptation, while the blue arrow is only performed in the case of interoception-based adaptation. Memory-based functions and goal-directed behaviors in the contextual layer are not elaborated. Reproduced from [2, 18, 20, 30, 31].

Figure 7 describes the inclusion of our proposed structure, following the World – Self – Action paradigm. Although the implementation of our allostatic control utilizes sensory input categorization, fosters autonomic alterations operating on a fairly prolonged temporal window, and goes beyond spontaneous reactions, it is not fully adaptive. This is due to the lack of a heuristic process of associative learning – which is considered to be a must-have

of the adaptive layer. Thus, we place the allostatic controller at the interface between the reactive and adaptive layers instead of being merely reactive as in most literature [2, 8, 31].

According to [32], exteroception is how the agent perceives its external “world”, i.e., the information about the relationship between the body and its surroundings. Exteroceptive information is acquired by the sensors that the robot possesses and include indicators such as the agent’s relative position to objects populating the environment (i.e., proximity), visual properties of those objects (e.g., colors, structures, etc.), auditory cues, haptic feedback, or other inertia measurements. On the other hand, interoception is a group of physiological signals derived from inside the body, relating to information about the state of well-being and energy, physical and chemical conditions, stress, and arousal levels. For example, the living system “feels” the internal variations of glucose level, pH, or body heat whenever exposed to extreme temperatures, consumes food, or exerts water. That kind of inner fluctuations – which we call “impact” – is converted and normalized by the homeostatic controllers in the second “self” column.

The “self” column additionally contains an allostasis controller. Both interoceptive and exteroceptive adaptation mechanisms will be implemented within this node. The allostatic controller receives both homeostatic signals and stored data from previous encounters with stimuli. Afterward, the information is fused and translated into newly proposed variables, e.g., accumulated allostatic load. The allostatic module gets access to, retrieves information, and makes modifications to an internal hierarchy that holds values of all weighting factors of all motivational drives. Drive intensities are consequently weighted and sent in parallel to the action selector in the third column. The action selector follows the basis of the mRF: it serves as a convergent bottleneck that transforms enormous information generated from the previous modules into internal representations of regulatory behaviors and makes the decision of which congruent action to trigger from those representations.

The action selector is authorized to access a stocked collection of behaviors in the “action” column, looks up and obtains specific modulating parameters of the selected behavior, and sends those parameters to a motor controller that configures and delivers control signals to the robot’s actuators, putting the demanded somatic movements and autonomic functions into practice. Although the behaviors encapsulated in our collection are still relatively poor, mediating the robot’s velocity by the strength of arousal level is the first step towards a richer behavioral repertory as explained in section 1.2.4.

2.2.2. Gradient and non-gradient-based subsystems

In order to validate the performance of our introduced model, we proceed to implement three typical homeostatic subsystems of the desert lizard, viz., thermoregulation, hunger,

and thirst. However, finding a reliable method to represent those systems computationally is not so trivial. Previous works in allostatic control have been done in various directions. Sanchez-Fibla et al. [8, 9] used vector fields, i.e, gradients, to compute the local perception of the robot, estimate the actual and desired values of the homeostatic subsystems, and control the robot’s navigation within the vector fields. Meanwhile, Vouloutsi et al. [2, 3] obtained homeostatic states and behavioral expressions without gradients. Both directions have their own advantages and limitations.

It appears that Sanchez-Fibla et al. concentrated their work on spatial navigation and low-cost path planning to a waypoint goal. Hence, it is reasonable for them to leverage the gradient method which has a reputation in this field. Moving along the gradient’s direction allows navigation to the local maximum/minimum location in the gradient, while moving perpendicularly establishes a safety parameter contour [33]. Both techniques are handy in the cases of retreating towards a refuge, flocking, or thermoregulation. Another advantage might be the degree to which the sensorimotor functions and spatial characteristics are coordinated, which leads to the faster identification of areas of interest.

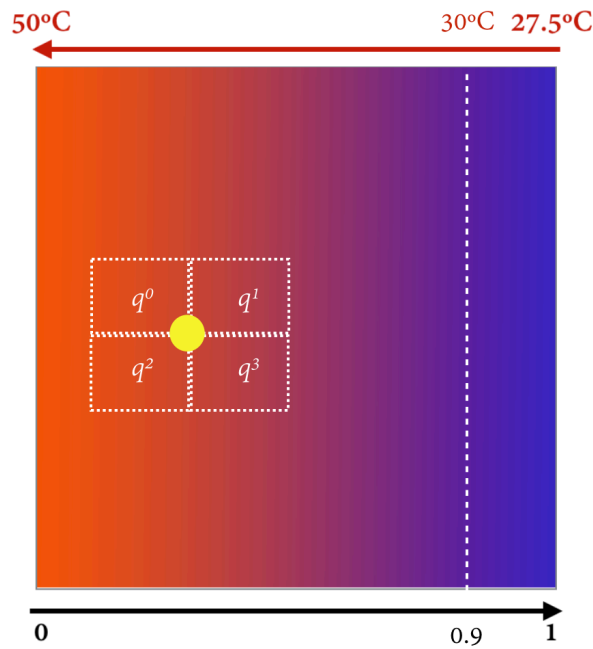


Figure 8: **The temperature gradient.** The temperature at the surface (left, orange) is at 50°C, and the robot is driven to move beneath the sand (right, blue) where the temperature is at 27.5°C–30°C. The yellow dot is the location of the robot, containing an actual temperature value and four quadrants of the local view.

We follow the gradient-based approach in [8] to build the temperature subsystem as a simple gradient composed of an actual value perceived by the agent’s local perception and multiple desired locations within the same vector field. The agent will try to place itself at an optimal distance to the sand surface, i.e., the highest peaks of a heat map to which the

gradient representing the temperature subsystem would correspond. Figure 8 depicts how we build the gradient as a 220x220 matrix to make it consistent with the experimental arena’s size. The ultimate left side of the matrix portraying the surface (with a temperature value of 50°C), and the right side represents cooler places beneath the sand. We inversely normalize the temperature from 50–27.5°C to a numeric value ranging from 0–1. The desired value is 0.9, approximately equivalent to 27.5°C–30°C. The x and y coordinates of the robot are interpreted as the row and column that make up a certain cell (containing a specific temperature value) in the matrix. It is worth mentioning that there is a constraint of local perception which means the agent has no knowledge of such a heat map. The local view comprises four quadrants (upper left, upper right, lower left, and lower right). Each quadrant includes the agent’s neighbor cells that form a smaller 4x3 matrix. q^0 , q^1 , q^2 , q^3 are the mean values of those four quadrants, and aV is calculated as the mean of the whole local view. Those quadrants are practically useful to steer the agent away from the so-called sand surface, with the h^{sign} controlling the increasing/decreasing direction of the gradient i and the AD^{sign} controlling the increasing/decreasing direction between the actual and desired values. Depending on how the actual value differs from the homeostatic limits, the system can ascend or descend the gradient to achieve constancy.

$$h_i^{sign} = \begin{cases} 1 & \text{if } q_i^0 < q_i^1 \\ -1 & \text{if } q_i^0 > q_i^1 \\ 0 & \text{otherwise} \end{cases} \quad AD_i^{sign} = \begin{cases} 1 & \text{if } aV_i < dV_i \\ -1 & \text{if } aV_i > dV_i \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

For instance, when the upper left quadrant q^0 is “warmer” than the upper right one q^1 , h^{sign} is set to 1, and the robot tends to turn right. In the opposite case, h^{sign} is set to -1, and the robot tends to turn left. Likewise, if dV is greater than the inverse aV , AD^{sign} is set to 1, and the robot is driven towards the cooler areas to reduce the actual temperature value. The AD^{sign} and h^{sign} values can be used to carry out the pertinent motor actions in order to bring the actual value closer to the desired one.

However, the agent in [8, 9] was provided with the local view of the constructed gradients instead of incoming data from the equipped sensors, going beyond constraints that animals in nature might have regarding their senses. By associating the homeostatic values with spatial locations of the gradient, the agent does not technically “feel” the fluctuations of internal variables, making the approach less reflexive and robust, particularly in the case of foraging. Vouloutsi et al. [2, 3] proposed more abstract techniques to put homeostasis and allostasis into practice through sensorimotor abstraction and knowledge representation. By allowing the input devices to coordinate and calibrate with each other, the researchers of [2] facilitated the mappings from sensor-centric information onto an ego-centric reference

frame that the robot used as its pivot perception. Once aligned, the raw sensory data is processed into symbols of interest with detailed properties. The control system persistently monitors the actual values of all homeostatic modules and activates alerts whenever there is a tendency to bring back the balance. The authors of [2] also provided each homeostatic subsystem with a natural decay factor that can compensate or detract it, which we replicate for the implementation of our hunger and thirst subsystems. The symbolic representations of the acquired environmental information are “target” and “impact” in our case. The “target” entity informs the agent the type of resources it is in contact with and its relative position to the resources (through a camera and two proximity sensors) – all are helpful for the robot to move towards or away from the salient stimulus. Meanwhile, the “impact” entity indicates fluctuations that such resources make to the related homeostatic subsystem after being consumed. For instance, the energy level of a living body will add up a certain amount whenever it consumes food, and resource impact refers to that incremental unit.

Thus, given a time step t and a decay factor b , depends on whether the robot is co-located with a relevant resource and receives an impact a , the actual value aV of the subsystem i will be updated as follows:

$$aV_i^{t+1} = \begin{cases} aV_i^t + a_i - b_i & \text{if } a_i > 0 \\ aV_i^t - b_i & \text{otherwise} \end{cases} \quad (2)$$

2.2.3. Interoceptive and exteroceptive cues for adaptation

In [8, 9], the weighting factor k is fixed and it simply controls the inertia suffered in each homeostatic subsystem, i.e., the weight of that subsystem. As explained in section 2.1, we argue that this k factor is inconstant and could rather be modified based on either the likelihood of finding a certain resource or the allostatic load.

In the first case of adaptation based on external sensing, the agent keeps track of all stimuli reached along its trajectory by storing them into an array sustained only during a specific time span. It is worthwhile to mention that the agent is constrained by its local sensing, and thus, the number of detected targets is not identical to the number of actual resources currently populating its habitat. There is a lack of a global view and resource identity here: the agent can identify only one object at a time, and it cannot tell whether the proximity sensors are sensing an object it came across before. This number of detected resources is inversely encoded as a numeric value ranging from 0–1. A higher encoded value represents a lower chance of encountering the desired stimuli. This value is close to 1 when the agent can barely spot any consumable target, and it approximates 0 when the

agent perceives an abundance of resources. In both strategies of adaptation, the weighting factor k is first initiated according to the hierarchical organization of drives described in section 1.2.1 (Figure 3). Given a time step t and the initiated value k^0 , below is the equation of how the weighting factor k of the behavior i is regulated by the number n of exteroceptive detected resources:

$$k_i^t = \begin{cases} k_i^0 \cdot (1 + \frac{1}{n_i^t}) & \text{if } n_i^t > 0 \\ 1 & \text{otherwise} \end{cases} \quad (3)$$

In the second approach of modulating k based on interoceptive cues, we aim at using the allostatic load as the mean value of tensions that need to be reduced in order to gain back homeostatic balance after a specific time period. “Drive intensity” refers to the absolute difference between the actual and desired values ($|aV - dV|$) of the homeostatic system i . During a predefined period, similarly to the exteroception-based direction, the agent stores all of the generated drive intensities into an array. We apply a low-pass filter to the array to pass signals with a frequency lower than a selected frequency of 0.2, and attenuate signals with frequencies higher than that cutoff frequency. Allostatic load is calculated as the mean value of that array, solving the requisite that the homeostatic system will relax if no drive intensity is generated after a while. In this respect, the k factor of the behavior i is adjusted by the allostatic load AL as the following equation:

$$k_i^t = k_i^0 \cdot (1 + AL_i^t) \quad (4)$$

In the particular case of temperature, since the thermoregulation subsystem possesses many privileges that might hamper other needs within a limited time of the simulation, we grant the agent with a higher tolerance of hotness by applying a discount of 0.2 to the temperature allostatic load. Those privileges include a higher initiated hierarchical order, a demanding desired value, the lack of a cyclical way for the actual value to decay gradually. All k factors of all motivational drives are subsequently integrated with drive intensities, and the final results – which we call “forces” – are important criteria for the action selector to determine the dominant motivational drive. The term “force” comes from the fact that we will apply this value to the motor translational speed at the very next time step.

$$f_i = 1 + k_i \cdot |aV_i - dV_i| \quad (5)$$

f integrates the current value of k and the difference between the actual and desired values ($|aV - dV|$) to weigh the contribution of a drive. Therefore, when two forces representing

weighted motivational tendencies are compared and both have the same value of intensity $|aV - dV|$, the dominant drive will be the one with the higher k . However, if the gradient i with a lower value of k_i has a higher intensity, the system might favor this i gradient.

2.2.4. Mechanism for action selection

The limited motor capacities of the mobile robot could be distributed to several behaviors running in parallel by gradient-based implementation. Besides computing the local views and the actual values of the gradients, Sanchez-Fibla et al. [8, 9] used the h^{sign} and AD^{sign} as modulating parameters bringing the aV closer to the dV to achieve homeostatic constancy. The agent specifies a certain strength to each gradient – ranging from no contribution at all to a maximal influence – by forming a vector. The magnitude of that vector is equal to the motivational tendency ($|aV - dV|$), while its angle indicates the ascending or descending gradient direction (controlled by h^{sign} and AD^{sign}) towards the desired location. The mapping to the left and the right wheels of the robot was done by summing up the contributions of all gradients. Thus, the selected behavior is the result of the combined motor contributions from all gradients, i.e., all motivational tendencies [8]. As a result, the lower-priority needs are not totally downplayed by the control system, and thus, Sanchez-Fibla’s model is more efficient in terms of sustaining a higher global stability. Since we apply both gradient-based and non-gradient-based techniques in our system, motor mappings in our study cannot be carried out that way. Our action selector ignores all other secondary needs, only chooses to trigger the behavior that satisfies the most dominant one. The formation of motor control commands in our study is also different from [8, 9] since we utilize the Robot Operating System (ROS) framework to build the software architecture instead. Our model has access to an internal collection of several modular behaviors, viz., obstacle avoidance, exploration, sand diving (moving towards the cool areas), and target reaching. This collection further holds all inherent motor parameters of the encapsulated behaviors. Those parameters can be applied to determine the robot’s translational speed v and rotational speed ω at the next time step. The system retains obstacle avoidance ($v = 0$, $\omega = -1$ for an obstacle on the left and $v = 0$, $\omega = 1$ for an obstacle on the right) and free exploration ($v = 1$ and a random ω ranging from 0–0.5) as the most and least important goals, respectively. For the sand diving behavior, we refer to the way Sanchez-Fibla et al. [8] calculated the forward and backward speeds of the robot utilizing h^{sign} and AD^{sign} , and put together two new equations for the translational and rotational speeds:

$$v_i^{t+1} = (1 + k_i^t \cdot |aV_i^t - dV_i^t|) \cdot v_i^0 \quad \omega_i^{t+1} = -1 \cdot h_i^{sign t} \cdot AD_i^{sign t} \cdot \omega^0 \quad (6)$$

For food and water approaching behavior, we apply the difference d between the current x coordinate of the agent and the current x coordinate of the detected resource to determine

the target rotational speed of the robot. This value d indicates whether the resource is on the agent's left or right side. If this value is positive, the resource is on the right side, and the robot turns right. If this value is negative, the resource is on the left side, and the robot turns left. The equation for the translational speed is similar to the sand diving behavior.

$$v_i^{t+1} = (1 + k_i^t \cdot |aV_i^t - dV_i^t|) \cdot v_i^0 \quad \omega_i^{t+1} = \begin{cases} 0.1 & \text{if } d_i^t > 0 \\ -0.1 & \text{otherwise} \end{cases} \quad (7)$$

2.3. Technical descriptions

In order to manipulate the robot and structure the interactions, we use the open-source ROS robotic framework [34]. ROS is not just a networking protocol but also an operating system in concept since it provides services such as hardware abstraction, low-level device control, common functionality, real-time rendered user interfaces, and messaging transport between modules. ROS also provides exploited sensors, ways to store and apply rotation matrices, plus convenient registration of new publisher and subscriber nodes.

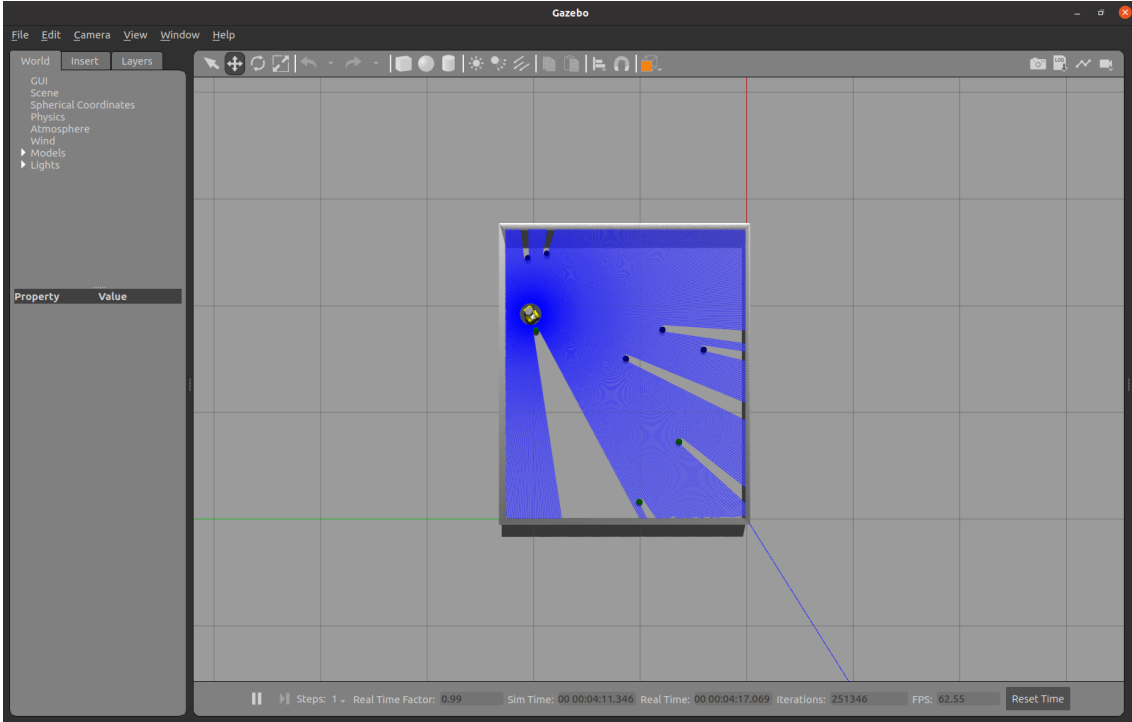


Figure 9: **Gazebo simulated environment.** The Gazebo world includes a rectangle experimental arena with wall boundaries. The green spheres populating the arena portray food, and blue ones represent water sources. The lower half of the arena bears a resemblance to the sand surface with high temperatures.

The overall system used in this study comprises the mobile robot simulated in a Gazebo environment (Figure 9) which we control using self-contained nodes written in Python. In

order to initiate a variety of behaviors, the simulated robot is allowed to send messages to the Python nodes and vice versa. For example, one ROS node can capture the images from a camera and send those images to another node for processing. Visual stimuli detection is achieved using image processing and feature extraction libraries such as the Open Source Computer Vision Library (OpenCV) [Bradski, 2000]. Various Python executable scripts are in charge of gathering data acquired from the Gazebo simulated environment, processing them into homeostatic and allostatic variables, and formulating roto-translation commands. Those control signals will be sent over via ROS protocol to the Gazebo simulator in order to allow vector-based movements to be executed.

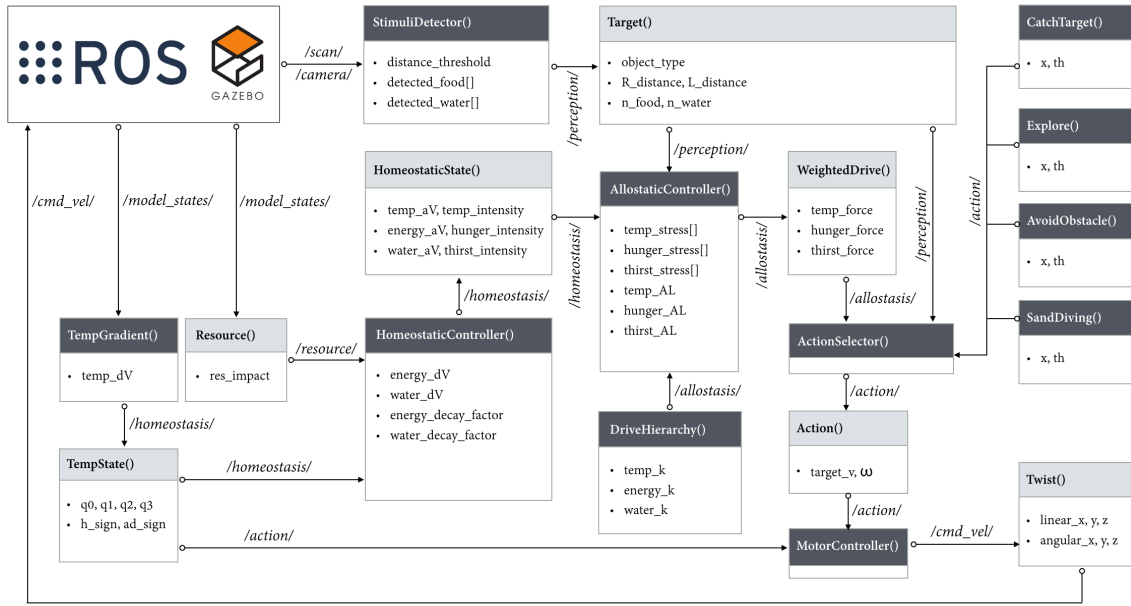


Figure 10: **Software system overview.** Messaging diagram between objects. Texts describe ROS topics and the direction of arrows indicate the direction of information delivery. Light grey objects represent messages, while dark grey objects resemble the components of the World – Self – Action framework.

Figure 10 explains in detail the mechanism used by ROS nodes to communicate through subscribers and publishers. ROS messages – which can be customized – are categorized into specific topics, and nodes may publish messages to a particular topic or subscribe to a topic to receive information. Following object-oriented programming principles, we decide to construct the components of our architecture and the custom messages as objects with properties as the elements of sensation, homeostasis, allostasis, and action. Objects might contain values that the agent acquires, processes, or generates when interacting with the Gazebo world, such as: sensory data, resource impacts, four quadrants of the local view, actual values, drive intensities, stresses and allostatic loads, weighting factors, or forces of all subsystems. Other objects hold inherent attributes of the model's components, e.g., the desired values, the decay factors, initiated values of the weighting factors, and the motor

parameters of stocked behaviors. The motor signal for a target vector-based movement is also formed as a messaging object.

2.4. Experimental design and setup

We implemented three different kinds of simulated agents to observe whether they can perform sufficient behavioral adaptations within the simulated Gazebo environment, and to what degree they can achieve organismic integrity. The probability to provide resources was randomly distributed, and we also expected to see the correlation between that and the agent's internal parameters. The Gazebo workspace consists of a rectangle 2.2x2.7m arena with four walls as boundaries. The synthetic two-wheeled mobile robot is equipped with a camera and two proximity sensors. It can freely navigate within the experimental arena, rotating in both directions: clockwise and counter-clockwise to head towards or to steer away from certain locations. Multiple homeostatic subsystems were implemented, viz., thermoregulation, hunger, and thirst. The temperature gradient is represented as a matrix associated with the experimental arena, being colder at one side and reaching the highest peak at the opposite side of the matrix. Water and food sources were allocated within the robot's explorable space as green and blue spheres, respectively. Blue spheres were distributed to one specific corner of the arena, and green spheres populated another area. All spheres were spawned to the environment and presented along the path of the agent, so the agent can autonomously execute an approaching or retreating behavior only when the corresponding homeostatic system is unbalanced. When the robot reaches the desired location and/or collocates with the target resource, the sphere is consumed (i.e., eliminated from the workspace), and the associated need is satisfied. The probability of the spherical resources was constantly modified during the agent's lifetime, starting with the minimum spawning speed of one spherical object per 2 seconds. The spawning speed adds up 2 seconds after every 5 seconds, and when it reaches the maximum threshold of one sphere per 50 seconds, the probability of resources decreases.

As for the setup of the synthetic agents, to ensure smooth simulations with explanatory results, we assigned and calibrated several arbitrary values that are inherent to the agent's internal systems and representational entities (Table 1). The hunger and thirst homeostatic systems possess three properties which demonstrate their characteristics in nature: the impact of relevant resources, the decay factor, and the minimum homeostatic limit. Living organisms tend to favor water over other nutritional sources since it is essential and needed in a greater degree than any other ingested substance, being classified as a macronutrient [21]. Thus, the agent's energy level – glucose level in this particular case – is proposed to have a lower decaying speed of the actual value (0.01 per second) in comparison to that of the water content (0.02 per second). Besides, every time the agent collocates with a “food”

sphere, its energy level increases with a unit amount (i.e., impact) of 0.2, while that value is 0.4 in the case of consuming water. The thirst homeostatic subsystem also has a more demanding desired value of 0.95 than that of the hunger one ($dV = 0.85$). The default ranking orders of homeostatic subsystems within the hierarchy of needs also differ. Hunger is the least prioritized motivational drive with initialized k of 0.78. The need for water has an initiated k value of 0.8, and the k value of the thermoregulation drive starts with the highest base of 0.85.

Table 1: **Arbitrary properties of homeostatic subsystems.** See text for further explanation.

Properties	Energy	Water	Temperature
Desired value dV	0.85	0.95	0.9
Resource impact a	0.2	0.4	N/A
Decay factor b	0.01	0.02	N/A
Initiated weighting factor k^0	0.78	0.8	0.85

The control system was built as per the descriptions in section 2.2. There are three versions of implemented agents. One operates under a simple reactive, autonomic mechanism with a static hierarchy of motivational tendencies. The weighting factors k were not adjusted in this version. The second version allows the agent to regulate the k factors adaptively based on exteroceptive sensing of the probability of environmental stimuli. The third version promotes an interoceptive agent configuring its hierarchy of internal drives based on the allostatic load. After every 1,000 time steps, the accumulated array of perceived stress (used to compute allostatic load) and the accumulated array of resources reached were cleaned, and both tracking processes reset. Each type of agent was tested during 10 simulations. Every simulation started with the same fixed position ($x = 0.1\text{m}, y = 1.1\text{m}$) of the robot, which means agents began their excursions at a cool location equidistant to both food and water sources. The motivation behind this setup is to have the very first observations of which adaptation mechanism might work best in helping agents sustain higher internal stability and ensure a reliable correlation between needs, actions, and resources.

Each simulation of our study lasted for approximately 15–20 minutes and included 500,000 time steps. At each time step, all the ROS nodes synchronously exchanged data, and those generated data was concurrently logged to CSV files which were used later for data analysis. The logged information from every single simulation includes: the trajectory of the agent, the number of resources detected, homeostatic information, allostatic parameters, and the dominant drive at every time step (Table 2). Data analyses for both individual simulations

and ones grouped by agent type were carried out, and the detailed results are explained in the following section 3.

Table 2: **Data collected from simulations.** Items marked with (*) are exclusively in the case of interoceptive agents, and items marked with (**) are collected for exteroceptive adaptation.

Collected data	Description
Time step	The current time step.
Robot's position	x, y coordinates of robot's location at every time steps.
Resources' position	Type (color) and x, y coordinates of resources during agent's lifetime.
Resource probability n (**)	The quantity numbers of food and water detected at every 1,000 time steps.
Desired values dV	The desire values of all homeostatic subsystems.
Temperature local view	The mean values of four quadrants q^0, q^1, q^2, q^3 of the temperature gradient at every time steps.
Temperature h^{sign} and AD^{sign}	The h^{sign} and AD^{sign} values of the temperature gradient at every time steps.
Actual values aV	The actual values of all homeostatic subsystems at every time steps.
Drive intensities $ aV - dV $	The absolute difference of the actual and desired values of all homeostatic subsystems at every time steps.
Allostatic loads AL (*)	The mean values of drive intensities of all homeostatic subsystems at every 1,000 time steps.
Weighting factor k	The weighting factors of all motivational drives at every time steps.
Force $f = 1 + k \cdot aV - dV $	The final forces (weighted drives) of all motivational drives at every time steps.

3. Results

3.1. Static and dynamic hierarchy of motivational drives

The first hypothesis to be proven in our study is whether the reconfigurable hierarchy of motivational tendencies holding dynamic k values can permit the agents to adapt better to a stochastic environment in comparison to a static hierarchical organization of drives with fixed k values. By analyzing both homeostatic and allostatic parameters generated during the simulations, we hope to prove this hypothesis statistically.

3.1.1. Results for homeostatic control

The primary objectives of behavioral regulation are to stabilize the physiological conditions within the living system and reduce the erroneous signals, i.e., the drive intensities to gain back homeostatic balance. Once all homeostatic systems are defined, the temporal series of actual values can be obtained during a tracking session equivalent to the agent's predefined lifetime of 500,000 time steps.

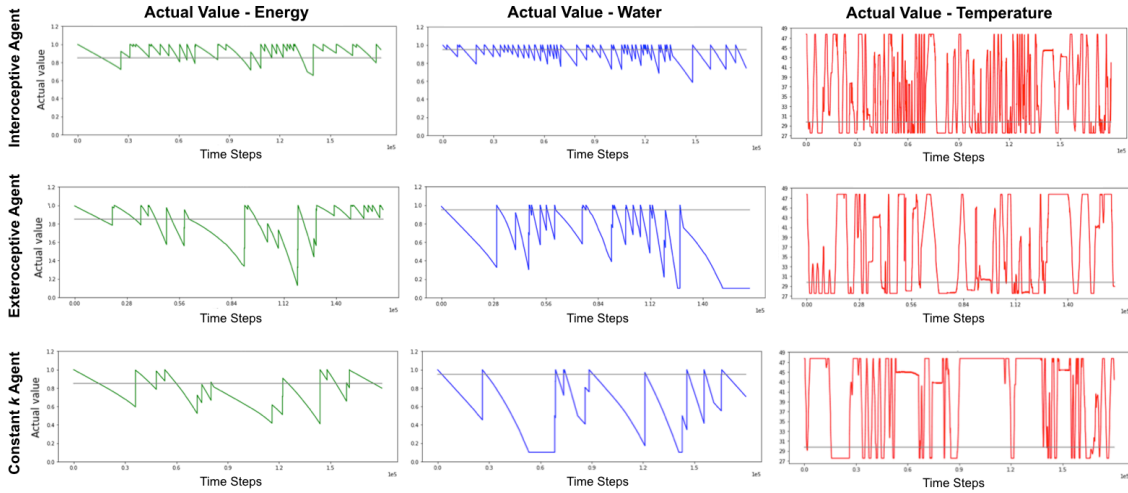


Figure 11: **Time series of individual actual values.** Each row of three plots comes from an individual simulation. The green plots indicate the actual energy level, the blue ones plot the actual water content, and the values constituting the red plots are extracted while the agent navigates within the temperature gradient. The grey lines bear a resemblance to the desired values of the corresponding homeostatic subsystems.

We chose three sample individual sessions; each comes from an agent performing one of the three particular ways to configure the weighting factor k (described in section 2.4). The actual values of the implemented subsystems were extracted at each time step and are shown in Figure 11. The hunger and thirst homeostatic systems suppose to keep the actual values higher than or equal to the desired values, while the thermoregulation subsystem should try to reduce the actual value lower than the favorable temperature of 30°C (higher than the inverted numeric aV). The organismic viability of an agent is considered to be

fatally endangered whenever the actual energy or water value reaches the bottom of the graphs ($aV = 0.1$), or when the temperature values stay atop at 50°C for a long while. The more the actual values stay within the permissible homeostatic ranges, the more likely the agent reaches the appropriate stimulus, the faster the homeostatic recovery process, and the better adaptations the agent performs.

By these time series charts, we can observe that the interoceptive and exteroceptive agents had the tendency to maintain the actual values of water and energy around the homeostatic minimum limit with fewer deviations (i.e., fewer outliers and drastic shifts caused by either the lack of detected resources or insufficient responses). Besides, it seems like they could react more quickly to return to the cooler areas of the temperature gradient, considering this time series as a collapse of the robot’s trajectory into one dimension [8]. Thus, we could see that the interoceptive and exteroceptive agents have the tendency to perform better adaptation and achieve better stability than the one with constant k . It is worthy to mention that due to a higher decay factor of water content, the energy system seemed to achieve better results in general. Likewise, the temperature system performed the worst since we did not implement an efficient way for the temperature value to decay gradually.

Table 3: **The mean and standard deviation values of actual values.** From all sessions of each agent type for each subsystem. Normality test p -value = 0.000 due to a large sample size of 5,000,000 time steps.

	Energy		Water		Temperature		
	<i>Mean</i>	<i>Std</i>	<i>Mean</i>	<i>Std</i>	<i>Mean</i>		<i>Std</i>
Interoceptive Agent	0.85	0.16	0.81	0.19	0.62	31.3	0.35
Exteroceptive Agent	0.83	0.15	0.74	0.22	0.60	32.1	0.36
Constant k Agent	0.76	0.19	0.66	0.25	0.56	33.3	0.36

Yet individual simulations cannot be generalized, and analyzing all simulations grouped by regulatory mechanisms might give us a more reliable conclusion. Data from all simulations of the same agent type were concatenated and both mean and standard deviation were computed. Table 3 depicts the mean and standard deviation values of the aV time series from all available sessions of each agent type for each homeostatic subsystem. The means support our hypothesis that both types of adaptive agents achieved higher viability than the one with fixed k . Also, the actual values measured from the interoceptive agents are slightly higher than those from the exteroceptive agents. However, we could not see any significant differences between the standard deviations, which means we cannot tell which kind of agent stabilized homeostasis with less variability. Since each homeostatic subsystem has a different desired value, we carried out the very same processes to achieve statistics results

for drive intensity ($|aV - dV|$). Table 4 gives us more precise judgments about how the agents were satisfying their needs or suffering from inadequate responses during the simulations. The results of the means and standard deviations of the drive intensities draw similar conclusions as those of the actual values, yielding no reliable implication.

Table 4: **The mean and standard deviation values of all drive intensities.** From all sessions of each agent type. Normality test p -value = 0.000 due to a large sample size of 5,000,000 time steps.

	Hunger		Thirst		Thermoregulation	
	<i>Mean</i>	<i>Std</i>	<i>Mean</i>	<i>Std</i>	<i>Mean</i>	<i>Std</i>
Interoceptive Agent	0.05	0.13	0.14	0.18	0.30	0.31
Exteroceptive Agent	0.06	0.11	0.21	0.21	0.33	0.33
Constant k Agent	0.13	0.16	0.29	0.25	0.36	0.33

We eliminated the temporal component to extract and plot information about which ranges of drive intensity that each agent group falls within the most. Figure 12 shows the bar histograms of the concatenated data from all simulations grouped by the type of agent. The drive intensity span was divided into three smaller ranges: $0 \leq 0.33 \leq 0.66 \leq 1$. The lower the drive intensities of a certain homeostatic system, the better the agent fulfills the relevant need, the more stabilized that internal system is. An effective adaptive mechanism should produce histograms of drive intensities with the bottom bars to be the longest.

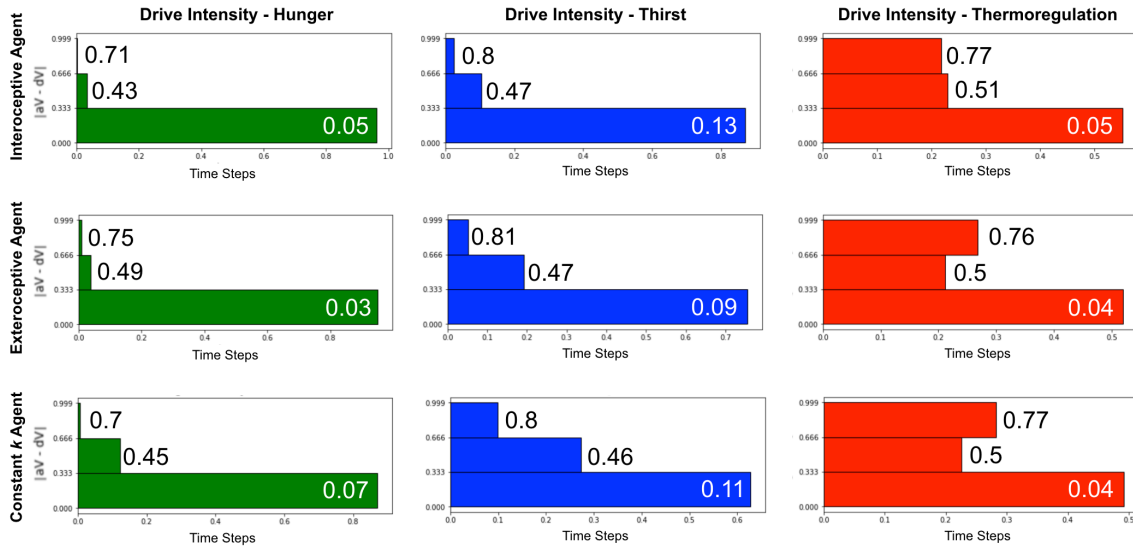


Figure 12. **Histograms of group drive intensities.** The $|aV - dV|$ span are separated in 3 smaller ranges: $0 \leq 0.33 \leq 0.66 \leq 1$. The number next to a bar indicates the mean value of that particular range.

We also ran Wilcoxon rank-sum and Levene's tests, getting p -values less than $7.2e-05$ when comparing all ranges of each histogram. Thus, we confirm that each histogram's 0–0.33

range differs from the other ranges (i.e., 0.33–0.66 and 0.66–1) significantly. With these histograms, we notice that the probability of generating drive intensities (i.e., homeostatic error signals) above 0.33 of the allostatic load-based agents is moderately lower than the exteroception-based agents in cases of all homeostatic systems. In other words, for each homeostatic subsystem, the ratio between the length of the bottom bar and the combined lengths of the top and middle bars of the interoceptive agent's is higher than that of the exteroceptive agent's (0.01, 0.2, 0.39 in comparison to 0.002, 0.1, 0.28 – for the hunger, thirst, and temperature subsystems, respectively). The results probably accentuate a higher potential to cancel internal perturbations out of allostatic load-based adaptation compared to exteroceptive-based adaptation.

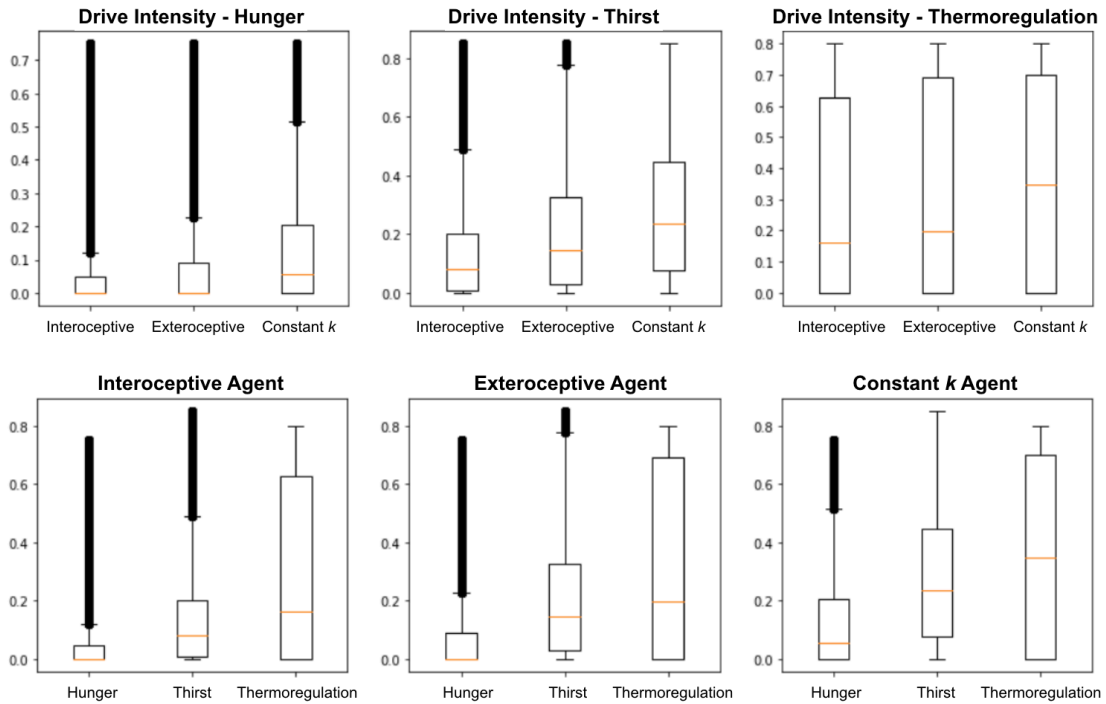


Figure 13. **Drive intensities grouped by homeostatic subsystems (top) and by agent types (bottom).** Orange lines indicate median values of drive intensity. The vertical length of boxes is the interquartile range.

We additionally expect to discover some promising advantages in behavioral adaptation of an agent calibrating the weighting factors based on allostatic load, making this approach equally efficient to modulating internal parameters directly from raw sensory data. We observe that the interoceptive agents appear to balance better the trade-offs between the three subsystems. For instance, in the case of adaptation based on external sensing of stimuli availability, the discrepancy between the mean value of thirst intensities and the mean value of hunger intensities is lower than that of the interoceptive agents and that of the agents with constant weighting factor k (0.09 in comparison to 0.15 and 0.16). To

examine this assumption, we finally proceeded to analyze and visualize the variability or dispersion patterns of drive intensities generated from each agent group in response to each homeostatic subsystem (Figure 13). The interoceptive agents generally created lower homeostatic error signals than the exteroceptive agents, and the exteroceptive agents also performed better than the agents with fixed k values, both in terms of lower median values and less dispersion of data. Although these outcomes do not reach statistical significance, we are convinced that longer simulations or better calibrated internal parameters might bring more reliable results supporting our hypotheses.

3.1.2. Results for allostatic control

Analyzing the actual values and drive intensities could only prove that the agents calibrating their hierarchy of motivational drives dynamically can achieve organismic integrity better than an agent with invariant weighting factors. However, whether the adaptation performed correlates with the constantly changing resource availability has not been analyzed yet. By investigating the temporal series of several internal variables (i.e., the number of sensed stimuli n , the allostatic load AL , the weighting factor k , and the force f – as explained in section 2.2.3), we can give more meaning to the behavioral regulation done by the agents.

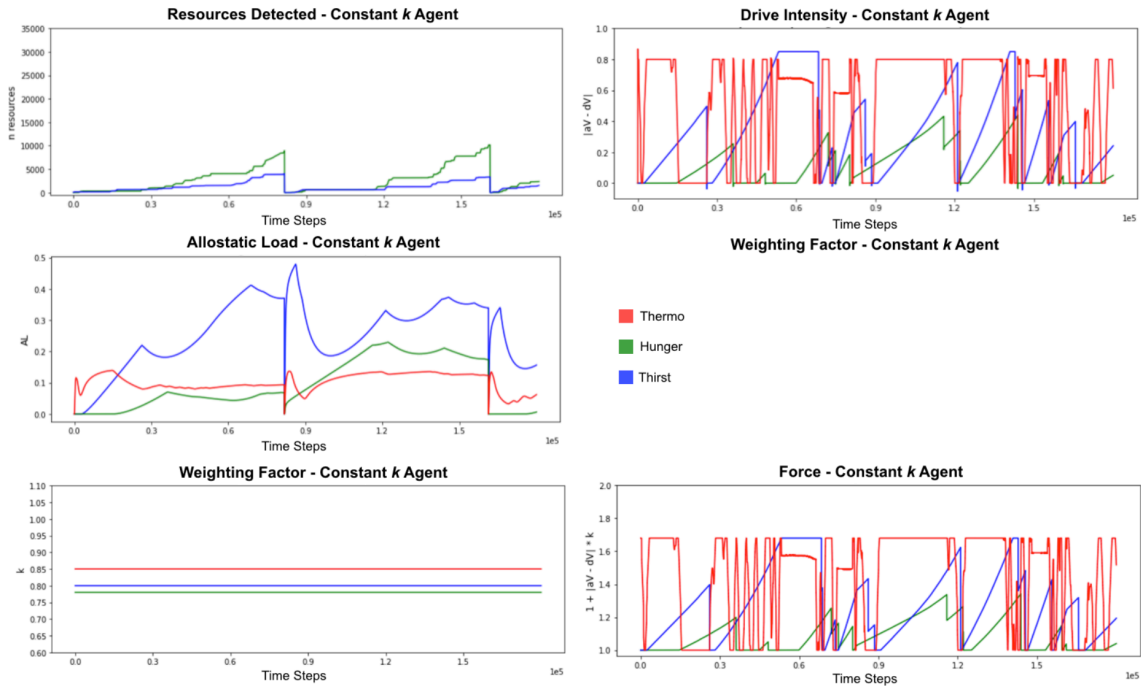


Figure 14. **Results for allostatic control of an agent with constant k .** The temperature gradient provides the agent with no visuospatial cues at all. The weighting factors remain the same during the simulation.

Behavior regulation can probably be executed without the alterations of the homeostatic subsystems' priority levels, but it would not be congruent with the availability of resources.

Figure 14 demonstrates the interrelation of the aforementioned variables occurring during a single simulation of an agent with a static hierarchical organization of needs. As shown in the plot of resources detected, a limited number of spheres were spawned to the arena and the agent was profoundly hungry and thirsty. Thus, allostatic loads caused by the scarcity of food and water sources were extremely high, correlating with the drive intensities of the related subsystems. However, due to the fixed priority levels applied uniformly through the session, the final force f at a given time step does not correspond significantly to the stress perceived by the agent at the moment. With such a shortage of water resources and low thermal stress, the agent was expected to perform the search for water more frequently and consistently. Instead, it just resembles the difference between the actual and desired values.

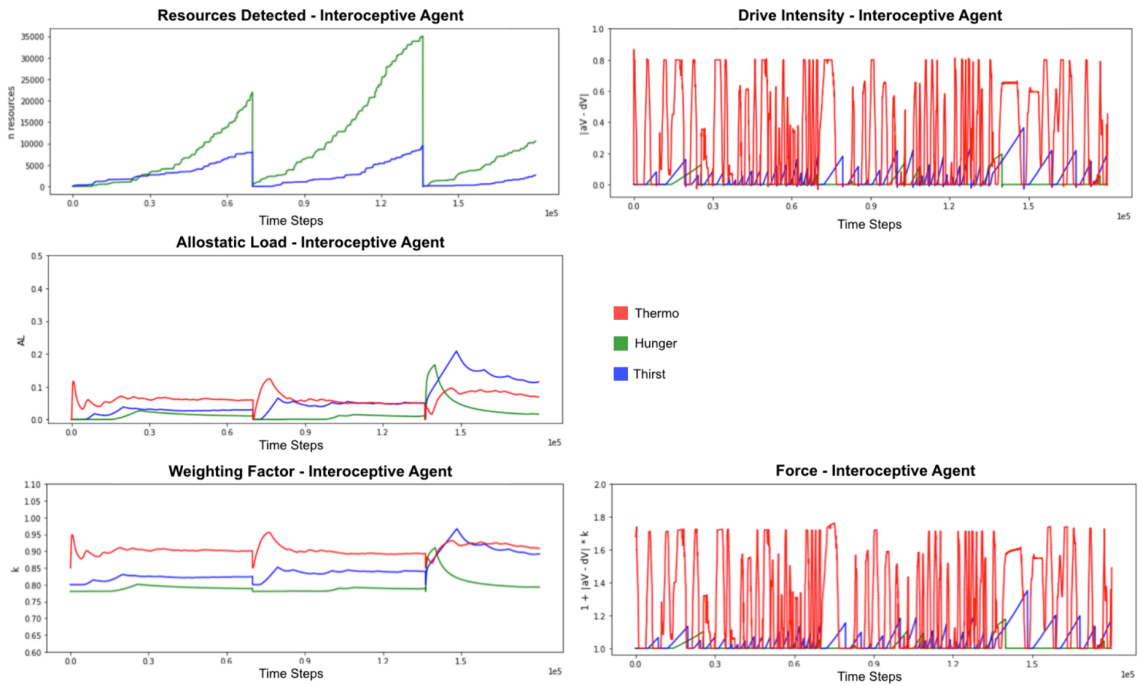


Figure 15. **Results for allostatic control of an interoceptive agent.** The weighting factors k correlate positively with the allostatic loads AL .

The interrelation of a single interoceptive agent's allostatic load AL and weighting factor k are explained in Figure 15. The agent could spot an abundance of food while the number of sensed water remained slow. The allostatic load at a certain time step obviously reflected well the concurrent drive intensity. Yet for the number of detected resources, their relation was not precisely expressed, especially at the second and last thirds of the simulation. The average discrepancy between identified food and water spheres during the second third is significantly huge in comparison to that of the last third, but the allostatic load – and thus, the weighting factor k – of thirst was escalated to a much greater extent during the last part of the session. The reconfiguration of the priority levels was performed fluidly during this

simulation with each motivational drive reached the top ranking at least once. However, the agent ventured onto the simulated sand surface several times, and thus, the need to move “beneath the sand” was too drastic. Therefore, we could not observe clearly how the force f was influenced by the k factor. Nevertheless, we aimed to properly and statistically calculate the correlation utilizing the concatenation of data from all sessions available.

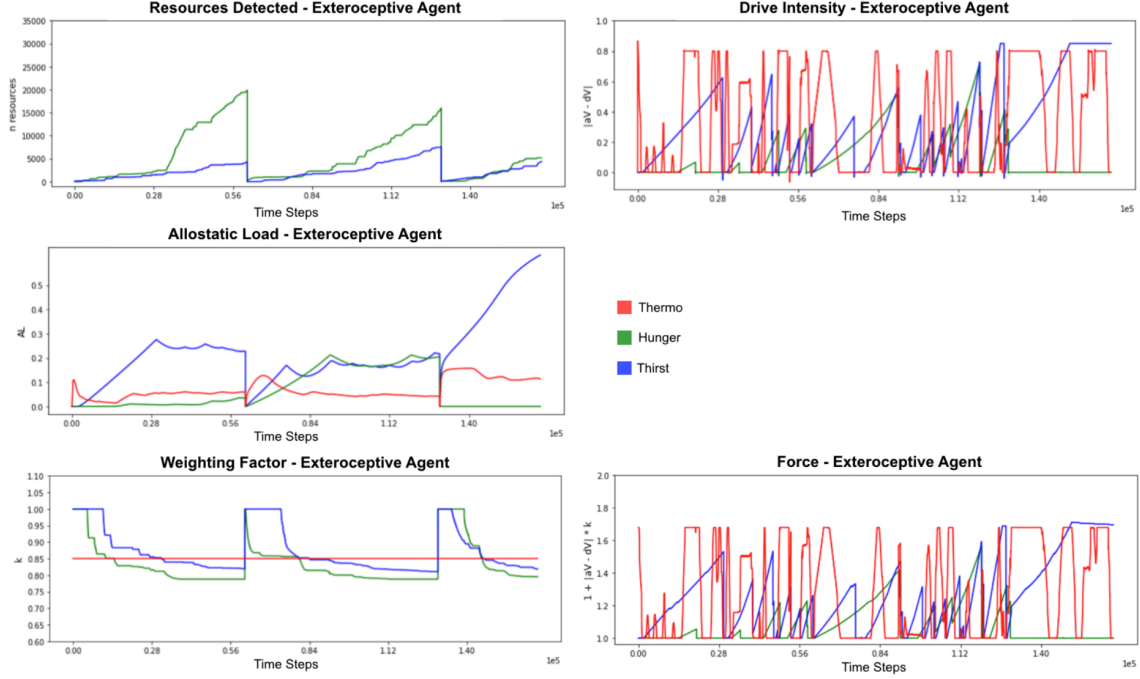


Figure 16. **Results for allostatic control of an exteroceptive agent.** The weighting factors k of food and water correlate negatively with the probabilities of detected resources n , whereas the weighting factor k of thermoregulation remains the same during the simulation.

Figure 16 shows similar charts for an individual exteroceptive agent. In contrast to the previous case of a stress-based adaptation strategy, the weighting factors k was modulated so that their values correlate negatively with the perceived probability n of environmental stimuli, ignoring what was happening inside the agent (i.e., signals such as drive intensities and allostatic loads). We noticed contradicting data during the first and last thirds of the chosen session. Despite the intense dehydration the agent was suffering from, the priority levels of the increased thirst needs during those time steps stayed low, subsequently deprioritized the tendency of the agent to select the action of drinking water.

3.2. Interoception and exteroception-based adaptive agents

The observations carried out during section 3.1 confidently support our first and second hypotheses about how a dynamic hierarchical organization of internal drives is an essential feature of sufficient regulation that ensures better organismic viability and integrity in both cases of adaptation strategies (described in section 2.2.3). To further support the second

hypothesis and validate the third one, we analyzed the trajectories of robots grouped by adaptation mechanism through logging their x, y coordinates within the experimental arena at every time step. We observe that both adaptive agents with the dynamic hierarchy of drives were able to prioritize navigations towards the food and water sources. At the same time, they tended to visit the right side of the arena (37.5°C–30°C) more frequently, while agents with a fixed hierarchy of motivational drives generate dispersed and unexplanatory trajectories that satisfy the most intense need for food, but fail to avoid the highest peaks of temperature. However, it should not be misunderstood that adaptive agents considered the secondary, less prioritized needs by optimizing locomotion towards resources following paths that ensure thermoregulation. That should be a future extension of our work.

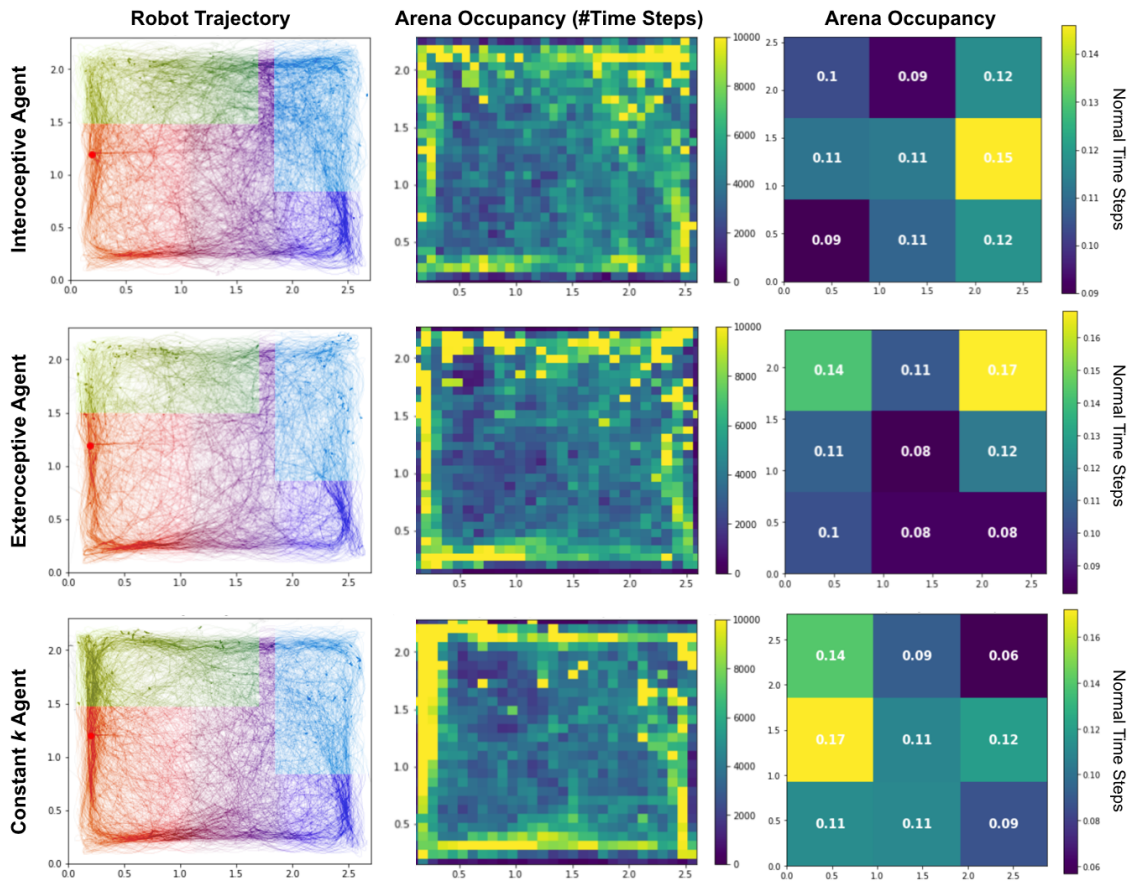


Figure 17. **Robot trajectories and arena occupancy maps.** Water is located upper right (light blue zone) and food is located upper left (light green zone). The left side of the arena is the sand surface with the peak temperature, and the other half represents cooler areas.

We additionally build occupancy 2D histograms and grid maps for each agent group. An occupancy grid is a discretization of space into fixed-sized cells, each of which contains a probability that it is occupied. The robots' concatenated trajectories and occupancy maps are depicted in Figure 17. The exteroceptive agents seemed to make optimal back and forth

excursions towards the specific corners where food and water spheres populated, probably because they were straightforwardly modulated by the external sensing of existing food and water resources. However, considering the occupancy at the cool locales, the allostatic load-based agents slightly ameliorated their regulated navigations. In addition, the interoceptive agents seemed to balance well the trade-offs between conflicting needs since efforts were allocated equally to all behaviors, which could be due to the fact that the weighting factor k of the thermoregulation drive is also flexibly adjusted.

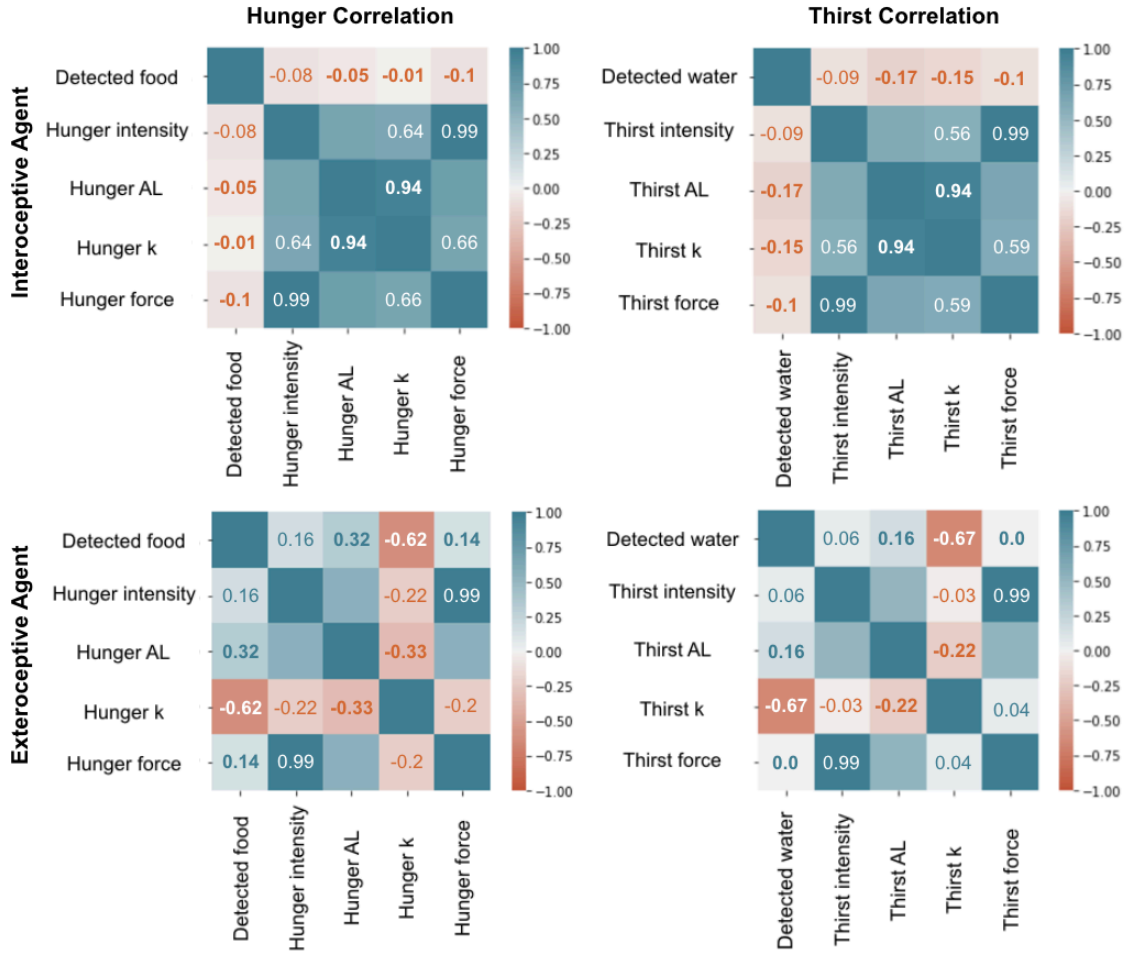


Figure 18. **Correlation matrices of internal parameters.** The numbers contained in cells are correlation coefficient r values.

A correlation analysis was done to quantify the relationship between internal parameters of the hunger and thirst homeostatic subsystems (Figure 18). With these correlation matrices, we found the exteroception-based strategy delivering several correlation coefficients that make no sense and just point out the irrelevancy of the weighting factor k in the reduction of internal drives. The oddest ones include the moderately negative correlations between the k factor and the drive intensity $|aV - dV|$ or the allostatic load AL . Besides, the k factor of the hunger subsystem and the force f applied to the food catching behavior are

negatively associated, which means the action selection was not influenced by k . Likewise, the positive correlation between the number of food detected and the hunger force proves insufficient adaptation to resource availability. This does not mean the way we stored the quantity of reached target is impractical, but future works should consider the triplets of state - action - reward instead of simply inverting the number of available resources into the value of the weighting factor k .

On the other hand, the interoceptive agents exhibited a slight tendency to have all internal variables negatively correlated with the number of detected resources n , especially the k factor and the force f . The results are not significant enough (i.e, the desired correlation coefficients of most relationships are lower than 0.3) for us to establish any evidence about how coherently the allostatic load-based agents adapted to the fluctuating probability of environmental stimuli, comparing to one based on external sensing of resources reached. Therefore, we cannot validate the third hypothesis but prove a potentiality of that.

Although we achieved a significant p -value = 0.000 due to the huge sample size of data, we acknowledge that many inherent properties of the homeostatic subsystems (enumerated in section 2.4) should be re-calibrated in order to yield remarkable results. Longer simulations or progressive techniques to tune the AL , k , and f variables could validate our suppositions about how allostatic load might play a key role in an intrinsic adaptive system.

4. Discussion

In this study, we have addressed the question of how behavioral self-regulation can be achieved through complementary homeostatic and allostatic principles. Allostasis plays a central role in the coordination of multiple homeostatic subsystems interacting together to withstand instability currently sensed by the living system. In addition, it is not obligatory for our allostatic controller to anticipate the organism’s internal states in prior. Instead, the congruous behavioral sequences emerge out of the reorganization of the priority levels of homeostatic subsystems. In some sense, we can consider allostasis as a meta layer defining the possibilities for the animal to adapt accordingly to fluctuating patterns of its habitat, being constrained by low-level cognitive processes.

We have implemented two adaptive versions of a simulated mobile robot that account for calibrating the importance levels of motivational tendencies based on exteroceptive and interoceptive environmental cues. The main principles of our computational model include decomposing behavior generation into functional modules incorporating physiological (i.e., allostatic load) and neuroscientific (i.e., the mRF) groundings that have not been explored fully in previous works regarding allostasis and action selection. We validated this model by testing thermoregulation and foraging behaviors of the robotic agents within a simulated Gazebo environment. We observed that the concept of a dynamic Maslow-like hierarchical organization of motivational drives where the priority levels of homeostatic subsystems are dynamically reconfigured can offer finer matchings between the agent’s internal needs and the opportunities or threats provided by its external surroundings.

As discussed in the literature review, little is known about how exactly allostatic control is realized in biological living systems. We proposed extensions to previous works in allostatic control, precisely the dynamic adjustments of the weighting factor k instead of altering the desired values dV of homeostatic subsystems [8]. Thus, we consequently pose the question concerning the possible coexistence and interplay between those two mechanisms, which could be a potential future extension of our implemented work. The allostatic adaptations to proximate situational triggers from constantly changing habitat configurations could be done through our proposed method. However, if the reason for adaptation derives from anticipatory disturbances and allostatic changes are going to persist for a longer temporal scale, changing the desired values dV would give rise to a more “evolutionary” perspective. Additionally, rather than the probability switching mechanism described in [8], future works can investigate how to adjust the homeostatic setpoints in another fashion. Examples could be modulating the dV in accordance with the impacts that autonomic systems can make on each other [21], or the predicted rates of change and time course of the new states [11].

As clarified in section 1.2.2, there is no predictive regulation carried out within the scopes of our study. Up to this point, the allostatic control described purely cope with immediate fluctuations that already occurred without forming any “prior knowledge”. Even though past encounters with external triggers or internal perturbations were used to follow a logic, the adaptive agents were totally unaware of upcoming events and lacked a heuristic to deal with environmental uncertainty. A broader self-regulation that allows contextual, memory-based processes and complex goal-directed behaviors together with allostatic control was inspected and evaluated in some preceding studies [31, 35, 36]. Future extensions should incorporate their findings and integrate our model into a complex cognitive architecture, with each layer concurrently functions on interfering purposes, i.e., the habitual reflexes handling proximate causes [35], or a goal-directed strategy allowing the agent to operate on longer time scales, utilizing information from the past (through memory) to make decisions appropriate for the future (i.e., planning) [36]. Subsequently, homeostatic error signals can be minimized either similarly to our direction of adjusting internal parameters, or via the constant updates of the higher layer’s predictions to approximate incoming sensory signals.

The adequate mediation between these layers can be solved with the higher layer being able to up-shift or down-shift its anticipations and regulates all low-level mechanisms in concert [37]. The mRF and the basal ganglia – which is traditionally considered to be in charge of action selection – were implied by [16] to form separate behavior generation mechanisms that interact with each other in such a combined hierarchical-layered manner. The basal ganglia’s primary route to the brainstem is through the pedunculo pontine nucleus, which itself projects heavily into the mRF [16]. Complicated computations are executed by the higher level (i.e., basal ganglia), and when all neuronal layers compute in parallel, the basal ganglia can veto the output of the mRF.

Some critiques can be made to our implemented structure, one is how it ignores secondary, less prioritized motivational tendencies. Our model is a simplification that works to solve drive conflicts without multi-purpose behaviors such as ensuring an optimal path towards the location of resources that avoids high temperatures. A potential direction could be the integration of multiple gradient maps that can elicit multi-system navigation. For instance, biological beings in nature can incorporate gradient-based subsystems such as geomagnetic fields, visual landscapes, the Coriolis forces, infra-sounds, and atmospheric chemosignals in order to form a single observable behavior [38]. The carrying through of internally defined and learned gradients influencing each other could underpin an integrative action selection process. Finding a plausible way to implement gradients resembling physiological measures (e.g., heart rate, skin conduction, etc.) will be valuable extensions for our control system.

5. List of figures

1.	Homeostatic control	8
2.	Detailed diagram of behavior generation	9
3.	The animal hierarchy of needs	10
4.	Dynamic configuration of the weighting factor	13
5.	Computational model of the medial Reticular Formation	15
6.	Behavior repertory for motivational tendency	16
7.	DAC's World – Self – Action framework	20
8.	The temperature gradient	22
9.	Gazebo simulated environment	27
10.	Software system overview	28
11.	Time series of individual actual values	32
12.	Histograms of group drive intensities	34
13.	Drive intensities grouped by homeostatic subsystems and by agent types	35
14.	Results for allostatic control of an agent with constant k	36
15.	Results for allostatic control of an interoceptive agent	37
16.	Results for allostatic control of an exteroceptive agent	38
17.	Robot trajectories and arena occupancy maps	39
18.	Correlation matrices of internal parameters	40

6. List of tables

1. Arbitrary properties of homeostatic subsystems	30
2. Data collected from simulations	31
3. The mean and standard deviation values of actual values	33
4. The mean and standard deviation values of all drive intensities	34

7. Bibliography

1. Hull, C. Principles of behavior: an introduction to behavior theory. *The Journal of Philosophy*. 40(20): 558 (1943).
2. Lallée, S., Vouloutsi, V., Munoz, M. B., Grechuta, K., Llobet, J. Y. P., Sarda, M. & Verschure, P. F. Towards the synthetic self: making others perceive me as an other. *Journal of Behavioral Robotics* (2015).
3. Vouloutsi, V., Lallée, S. & Verschure, P. F. Modulating behaviors using allostatic control. *Conference on Biomimetic and Biohybrid Systems*. 287–298 (2013).
4. McEwen, B. S. & Wingfield, J. C. The concept of allostasis in biology and biomedicine. *Hormones and behavior*. 43: 2–15 (2003).
5. Bradshaw, S. Homeostasis in desert reptiles. *Adaptations of Desert Organisms* (1997).
6. Houston, A. & McFarland, D. On the measurement of motivational variables. *Animal Behavior*. 24: 459–475 (1976).
7. Sterling, P. & Eyer, J. Allostasis: A new paradigm to explain arousal pathology. *Handbook of life stress, cognition, and health* (1988).
8. Sanchez-Fibla, M., Bernardet, U., Wasserman, E., Pelc, T., Mintz, M., Jackson, J., Lansink, C., Pennartz, C. & Verschure, P. F. Allostatic control for robot behavior regulation: A comparative rodent-robot study. *Advances in Complex Systems*. 13(3): 377–403 (2010).
9. Sanchez-Fibla, M., Bernardet, U. & Verschure, P. F. Allostatic control for robot behavior regulation: An extension to path planning. *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 1935–1942 (2010).
10. Lee, D. Y., Kim, E. & Choi, M. H. Technical and clinical aspects of cortisol as a biochemical marker of chronic stress. *BMB Reports*. 48(4): 209–216 (2015).
11. Sterling, P. Allostasis: a model of predictive regulation. *Physiology & Behavior*. 106(1): 5–15 (2012).
12. Maslow, A. A theory of human motivation. *Psychological review*. 50(4): 370–396 (1943).
13. Guerrero-Rosado, O. & Verschure, P. F. Robot regulatory behavior based on fundamental homeostatic and allostatic principles. *Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence* (2020).
14. Blouet, C. & Schwartz, G. J. Hypothalamic nutrient sensing in the control of energy homeostasis. *Behavioral Brain Research*. 209(1): 1–12 (2009).

15. Calderon, D., Kilinc, M., Maritan, A., Banavar, J. & Pfaff, D. Generalized CNS arousal: An elementary force within the vertebrate nervous system. *Neuroscience and Bio-behavioral Reviews*. 68: 167–176 (2016).
16. Humphries, M. D., Gurney, K. & Prescott, T. J. Is there a brainstem substrate for action selection?. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 362(1485): 1627–1639 (2007).
17. Humphries, M. D., Gurney, K. & Prescott, T. J. The brainstem reticular formation is a small-world, not scale-free, network. *Royal Society B: Bio-logical Sciences*. 273(1585): 503–511 (2006).
18. Humphries, M. D., Gurney, K. & Prescott, T. J. Is there an integrative center in the vertebrate brain-stem? A robotic evaluation of a model of the reticular formation viewed as an action selection device. *Adaptive Behavior*. 13(2): 97–113 (2005).
19. Breazeal, C. A motivational system for regulating human-robot interaction. *National Conference on Artificial Intelligence*. 54–61 (1998).
20. Kenrick, D. T., Griskevicius, V., Neuberg, S. L. & Schaller, M. Renovating the pyramid of needs: Contemporary extensions built upon ancient foundations. *Perspectives on psychological science*. 5(3): 292–314 (2010).
21. Breathnach, C. S. The hierarchy in homeostasis. *Irish Journal of Medical Science*. 140: 183–188 (1971).
22. Tschantz, A., Barca, L., Maisto, D., Buckley, C. L., Seth, A. & Pezzulo, G. Simulating homeostatic, allostatic and goal-directed forms of interoceptive control using Active Inference. *bioRxiv Preprint* (2021).
23. Corcoran, A. W. & Hohwy, J. Allostasis, interoception, and the free energy principle: Feeling our way forward. *The interoceptive mind: From homeostasis to awareness*. 15: 272–292 (2018).
24. Wingfield, J. C. The concept of allostasis: Coping with a capricious environment. *Journal of Mammalogy*. 86(2): 248–254 (2005).
25. Wingfield, J. C. Control of behavioral strategies for capricious environments. *Animal Behaviour*. 66(5): 807–816 (2003).
26. McEwen, B. S. Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiological Reviews*. 87(3): 873–904 (2007).
27. Cooper, W. & Pérez-Mellado, V. Tradeoffs between escape behavior and foraging opportunity by the balearic lizard. *Herpetologica*. 60(3): 321–324 (2004).

28. Martins, M. Defensive tactics in lizards and snakes: The potential contribution of the Neotropical fauna. *Anais de Etologia*. 14: 185–199 (1996).
29. Merker, B. Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences* (2007).
30. Verschure, P. F. & Duff, A. Distributed Adaptive Control: Theory and Practice. *Biologically Inspired Cognitive Architectures*. 1: 55–72 (2012).
31. Marcos, E., Sanchez-Fibla, M. & Verschure, P. F. The complementary roles of allostatic and contextual control systems in foraging tasks. *Simulating Animal Behavior (SAB) LNCS Springer-Verlag* (2010).
32. Craig, A. D. How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience*. 3(8): 655–666 (2002).
33. Adamek, T., Mas, I. & Kitts, C. A. Gradient-based cluster space navigation for autonomous surface vessels. *IEEE/ASME Transactions on Mechatronics*. 20(2): 506–518 (2014).
34. Kay, J. Proposal for Implementation of Real-time Systems in ROS 2. *ROS.org* (2016).
35. Maffei, G., Herreros, I., Sanchez-Fibla, M., Friston, K. J. & Verschure, P. F. The perceptual shaping of anticipatory actions. *Proceedings of the Royal Society B: Biological Sciences*, 284 (2017).
36. Freire, I. T., Amil, A. F., Vouloutsi, V. & Verschure, P. F. Towards sample-efficient policy learning with DAC-ML. *Procedia Computer Science*, 190, 256–262 (2021).
37. Freire, I. T., Moulin-Frier, C., Sanchez-Fibla, M., Arsiwalla, X. D. & Verschure, P. F. Modeling the formation of social conventions from embodied real-time interactions. *PloS one*, 15 (2020).
38. Wallraff, G. Avianavigation: Pigeon homing as a paradigm. *Springer-Verlag* (2005).