ZFN — Philosophical Problems in Science

ARTICLE

Homeostasis as a foundation for adaptive and emotional artificial intelligence

Anna Sarosiek

University of the National Education Commission, Kraków
Email: anna.sarosiek@uken.krakow.pl

Abstract

Homeostasis, a fundamental biological mechanism, enables living organisms to maintain internal balance despite changing environmental conditions. Inspired by these adaptive processes, research into artificial intelligence (AI) seeks to develop systems capable of dynamic adaptation, introspection, and empathetic interactions with users. This article explores the potential of implementing homeostatic mechanisms in AI as a foundation for emotional intelligence and self-regulation. Key questions include the distinction between simulation and actual experience, the role of machine introspection, and the emergence of qualitative states akin to phenomenal experiences. Drawing on Antonio Damasio's theory and classical concepts from cybernetics, the article investigates how homeostatic principles might inspire the development of AI, paving the way for more flexible and context-aware technologies.

Keywords: homeostasis, artificial intelligence, adaptive systems, empathy, self-regulation, introspection, machine introspection, Antonio Damasio, emotional intelligence, cybernetics

Introduction

Homeostasis, a fundamental biological mechanism, enables living organisms to maintain internal equilibrium in the face of dynamic changes in their environment. This feedback-based process allows organisms not only to respond to disturbances but also to anticipate future challenges, making it a cornerstone of adaptability. These regulatory and adaptive capacities have inspired researchers and engineers to explore ways of modeling such processes in artificial intelligence (AI). The natural homeostatic processes observed in living organisms can serve as a template for developing AI systems capable of autonomous operation, dynamic adaptation, and efficient resource management under changing operational conditions.

Modern AI systems, while advanced, remain constrained by their focus on optimizing performance within predefined goals. Introducing mechanisms inspired by homeostasis offers new possibilities, enabling systems to achieve mechanical introspection, dynamically balance internal and external disturbances, and engage in empathetic interactions. A key question emerges: how can homeostatic mechanisms inspire AI to move beyond simple reactivity to stimuli, toward decision-making based on internal and external signals?

This article explores the role of homeostasis in biology and its application to the development of artificial intelligence. Drawing on the foundational work of cybernetics pioneers such as Norbert Wiener and William Ross Ashby, as well as contemporary theories from Antonio Damasio, it argues that homeostasis is not merely a regulatory process but also the foundation of emotional and decision-making processes. Finally, the article examines the potential for leveraging these mechanisms in adaptive AI systems that could form the basis of future empathetic technologies.

Homeostasis in Biology

Homeostasis is a biological phenomenon that refers to the ability of living organisms to maintain internal equilibrium despite changing external conditions. The term was introduced in the 19th century by Claude Bernard, who observed that survival requires the constant regulation of internal functions, irrespective of environmental dynamics (Fleming 1984). Homeostasis, therefore, refers to maintaining stable parameters. For mammals, these include body temperature, blood glucose levels, pH, and blood pressure—all of which are critical for the organism's functioning.

Homeostatic regulatory mechanisms rely on feedback systems—primarily negative feedback—which allow organisms to respond to environmental changes and restore balance. Negative feedback acts as a control system in which the organism's response counteracts the disturbance (Perrimon and McMahon 1999; Bielecki 2016; Hancock et al. 2017). For example, in response to an increase in body temperature, sweating is activated to lower the temperature to an optimal level. While negative feedback is the dominant homeostatic mechanism, positive feedback also occurs in organisms, particularly in processes that require rapid responses (Peters, Conrad, and Hubold 2007). An example is blood clotting, where the activation of one clotting factor accelerates the activation of others until the vessel damage is sealed. Thus, homeostasis is not a static state but a dynamic process of constant adjustments that the organism makes to maintain stability.

In the 20th century, Walter Cannon significantly expanded the concept of homeostasis, describing it as the organism's ability to maintain internal stability in the face of external changes, which he termed "dynamic equilibrium." Cannon paid particular attention to the organism's responses to environmental stressors, such as temperature changes or the presence of pathogens, and introduced the term "fight or flight" to describe the automatic response to threats, characterized by increased heart rate, heightened adrenaline secretion, and energy mobilization (Modell et al. 2015). His research demonstrated how homeostatic mechanisms enable organisms not only to react to disturbances but also to prepare for future challenges.

Homeostasis in Cybernetics

The The concept of homeostasis has played a fundamental role not only in biology but also in the development of cybernetics—a discipline pioneered in the mid-20th century by Norbert Wiener and William Ross Ashby. Cybernetics, described by Wiener as the science of "control and communication in the animal and the machine," provided groundbreaking frameworks for understanding how systems maintain equilibrium through feedback loops (Wiener 1965). This interdisciplinary approach combined biological principles with machine

design, exploring how self-regulation mechanisms could enable both natural and artificial systems to dynamically adapt to changing environments.

A key contribution to cybernetics was William Ross Ashby's book *Design for a Brain (Ashby, 1952)*, which established foundational principles for applying homeostatic concepts to machine intelligence. Ashby's theoretical and experimental work introduced the "homeostat"—a device capable of maintaining stability by dynamically adjusting its parameters in response to external disturbances (Ashby 1952). The homeostat demonstrated the potential for machines to autonomously achieve adaptive stability, serving as a precursor to modern adaptive systems. Ashby's notion of "essential variables," critical for a system's functionality, mirrored the biological understanding of critical parameters like body temperature or blood pressure in living organisms.

In addition to Ashby's work, Grey Walter's cybernetic "tortoises" offered a compelling demonstration of how simple feedback mechanisms could generate seemingly complex behaviors. Developed in the 1940s, these autonomous robots were equipped with basic sensors that allowed them to seek out light sources and avoid obstacles, effectively simulating primitive forms of adaptation and goal-directed behavior (Walter 1950). The tortoises' ability to navigate their environment illustrated how feedback-driven processes could emulate biological homeostasis, blurring the line between natural and artificial systems.

These early examples of cybernetics not only highlighted the adaptability enabled by feedback mechanisms but also underscored the fundamental importance of homeostatic principles in the pursuit of machine intelligence. By showcasing how regulation and dynamic adjustment could be mechanized, Ashby and Walter laid the groundwork for modern efforts to develop artificial intelligence systems capable of empathy, introspection, and emotional intelligence. Integrating these concepts into contemporary AI aligns with the vision of adaptive, self-regulating systems that reflect the resilience and flexibility of biological organisms.

Homeostasis in AI: Beyond Feedback Loops and Evolutionary Stability

Modern artificial intelligence systems, such as neural networks, already employ feedback mechanisms in processes like backpropagation. However, these systems are primarily focused on optimizing performance based on external inputs and predefined goals. The concept of homeostasis, as presented in this article, introduces a higher level of complexity, emphasizing the potential for machine introspection—the ability of an AI system to analyze and adapt its internal processes in response to both internal and external disturbances. Unlike traditional feedback mechanisms, which adjust parameters to optimize specific tasks, homeostatic AI would integrate dynamic self-reflection and self-regulation, essential for achieving emotional intelligence and adaptive responses that emulate the resilience and flexibility observed in biological systems (Damasio and Carvalho 2013; Gros 2021).

While homeostasis and evolutionary stable strategies (ESS) both aim for stability, their mechanisms and assumptions differ significantly. ESS focuses on strategies that remain stable in competitive conditions, emphasizing external optimization, whereas homeostasis relies on dynamic feedback to maintain internal equilibrium and functionality in changing environments (Maynard Smith 1974; Dawkins 1989). For instance, ESS explains why certain strategies dominate within a population, but homeostasis provides a framework for

understanding how an AI system dynamically maintains balance by reconciling internal and external demands (Walter 1950; Ashby 1952).

In the context of AI, homeostatic mechanisms would enable systems to balance external pressures, such as user demands or environmental changes, with internal adjustments, such as resolving computational conflicts or managing resource constraints. This capacity aligns with concepts of neuroplasticity, where dynamic restructuring allows systems to adapt more effectively over time (Zhang, Shen, and Sun 2022). Machine introspection in AI could encompass the system's ability to evaluate its internal states, including resource levels, operational efficiency, or emotional responses to external stimuli. Such an approach connects to Searle's "Chinese Room" argument, in which AI not only simulates understanding but also considers the nature of its cognitive processes (Searle 1980). Introducing introspection could enable AI to dynamically modify its algorithms in real-time, fostering more advanced forms of adaptation.

By moving beyond reactive optimization, homeostatic AI offers transformative potential, integrating mechanisms of self-regulation and resilience. This progression paves the way for systems that mimic the dynamic equilibrium seen in biological organisms, opening avenues for adaptive, context-sensitive, and introspective AI technologies (Tegmark 2018).

The Extended Theory of Homeostasis

In the perspective of Antonio and Hanna Damasio, homeostasis is presented even more broadly. It is not only a biological mechanism but also a profound foundation for all emotional, decision-making, and social processes. According to them, homeostasis encompasses not only physiological processes but also psychological mechanisms that lead to the creation of emotions, feelings, and values (Korn and Bach 2015; Damasio and Damasio 2022). Thus, the concept of homeostasis should be expanded to recognize it as a driving force not only for individual survival but also for the development of more complex forms of mental and cultural life.

In Damasio's concept, emotions serve as homeostatic signals that integrate physiology with decision-making processes (Damasio and Carvalho 2013). Emotions, in his view, are responses to homeostatic disruptions and serve an adaptive function. For example, experiencing fear in a dangerous situation is not merely an emotional reaction but also a mechanism that mobilizes the organism to act and protect itself from harm. This framework suggests that artificial intelligence systems could integrate homeostatic signals analogous to emotions, enabling dynamic decision-making that mirrors biological adaptability. By interpreting external stimuli as disruptions to an internal equilibrium, such systems could prioritize actions not merely based on task efficiency, but on maintaining their own operational stability and responsiveness to human emotional contexts.

Emotions, as an integral part of homeostatic processes, enhance the chances of survival and improve quality of life. As a result, decisions made by individuals are deeply rooted in their physiological and emotional states and are thus linked to the fundamental drive to maintain homeostasis (Burdakov 2019). Bechara and Damasio also argue that human choices—from everyday decisions to more complex ones—are shaped by how a given situation affects internal balance, suggesting that homeostasis is a hidden mechanism regulating not only emotional reactions but also rational cognitive processes (Bechara, Damasio, and Damasio 2000). The

integration of homeostatic principles into AI could also foster introspection, allowing systems to evaluate and adjust their internal states, much like how emotions enable organisms to reconcile physiological and psychological needs.

From Damasio's perspective, homeostasis is also a deeply integrated system influencing cultural development (Kłóś 2014; Damasio 2018). On a social level, mechanisms that originally served individual adaptive functions have evolved to support human interactions, which was crucial for the development of complex societies. Damasio argues that the need to maintain physiological and psychological balance is the driving force behind the creation of social norms, rituals, morality, and cultural structures (Damasio 2018). For instance, empathy and compassion can be understood as mechanisms enabling social support for others, thereby supporting the shared homeostasis of the community. Unlike Evolutionary Stable Strategies (ESS), which are defined by static stability under competitive pressures, homeostatic mechanisms emphasize dynamic adaptability, making them better suited for real-time responses in variable and unpredictable environments.

This broader interpretation of homeostasis, encompassing physiological, psychological, and social dimensions, provides a valuable blueprint for developing adaptive and empathetic AI systems capable of navigating complex human contexts.

Homeostasis and Emotional Adaptation

In biology, homeostasis serves as the "guardian" of an organism's stability, regulating key parameters in response to external stimuli (Giordano 2013; Davies 2016). Antonio Damasio's interpretation of this concept extends its significance to emotions, decision-making, and social interactions (Damasio and Damasio 2022). Emotions provide individuals with essential information about which actions promote internal equilibrium, serving as both warning and adaptive functions. Fear, as a reaction to threat, mobilizes the organism to act, while joy signals favorable conditions (Pio-Lopez, Ramirez, and Santos 2023). These emotional states, which result from homeostatic regulation, help organisms choose appropriate strategies depending on the situation (Goldstein 2019).

Homeostasis enables biological organisms to respond to immediate disruptions as well as adapt based on past experiences. These processes form the basis of intelligence, as they allow for the development of more sophisticated adaptive mechanisms (Torday 2015) Moving beyond physiological and emotional responses to the level of conscious decision-making is an advanced form of homeostasis. Decisions that support survival and long-term benefits represent a form of organism regulation in a more complex environment. Biological intelligence allows organisms to predict potential disruptions and prevent them before they occur. This is possible through homeostatic mechanisms that "anticipate" environmental changes based on previous patterns (Eskov et al. 2017).

As organisms become more complex, their adaptive mechanisms evolve to include cognitive processes, allowing for more intricate information processing and more effective management of equilibrium in dynamic conditions (Davies 2016) Thus, we speak of the ability to interpret signals from internal homeostatic processes and consciously use them for decision-making (Billman 2013). Homeostasis lays the foundation for more advanced information processing, encompassing both emotional reactions and conscious decision-making. In other words, homeostasis is a fundamental "predictive mechanism"—it allows an organism

to react and adapt not only to immediate stimuli but also to future challenges. It is this ability to predict and manage internal balance that forms the foundation of emotional intelligence. At this stage, emotional intelligence emerges as the capacity to learn, make complex decisions, and solve problems in order to maintain internal equilibrium.

If homeostasis is the foundation of emotional intelligence in biological organisms, it is worth considering how homeostatic mechanisms could be modeled in artificial systems. Could introducing such mechanisms contribute to the development of empathetic artificial intelligence?

Homeostatic Mechanisms as an Inspiration for AI

Homeostasis provides inspiration for understanding intelligence in terms of adaptive and emotional capabilities. Introducing the concept of homeostasis into artificial intelligence could support AI's ability to respond to unforeseen situations, adapt quickly, and understand and interpret emotions in human interactions (Gros 2021; Zhou 2021).

Theories of adaptation in biology assume that organisms evolve by developing mechanisms that enable them to effectively respond to environmental challenges. Emotions play a key role in this process by providing quick and intuitive reactions to external stimuli. Translating these assumptions into artificial intelligence allows for a new understanding of adaptive intelligence as a dynamic process in which emotions are an integral part of effective response to changing environmental conditions (Assunção, Silva, and Ramos 2022; Zhao, Simmons, and Admoni 2022). Adaptation, where emotional equilibrium mechanisms become a dynamic reference point, would allow AI to develop more complex and contextual decision-making abilities. Introducing emotions as an adaptive factor in artificial intelligence expands AI's capacity to respond to a changing environment in a more flexible and predictable manner. This approach could lead to the development of AI that not only processes data on a cognitive level but also uses "emotional reference points" to assess situations and choose optimal action strategies.

The aforementioned "emotional reference points" could find practical applications through mechanisms of machine introspection. Introspection would enable AI not only to analyze its internal states, such as resource levels or emotional responses, but also to dynamically adjust its action strategies in real-time. Through introspection, a system could better interpret the emotional contexts of the user, transforming references into concrete operational decisions that support empathetic and flexible reactions. Such capabilities would be particularly useful in dynamic and complex situations, where the ability to analyze AI's internal processes and adapt them could significantly improve interactions with users and the system's capacity to respond to unforeseen circumstances.

Biological organisms have developed various survival mechanisms, including physiological regulation, as well as complex cognitive and emotional mechanisms. Emotions such as fear, joy, or empathy play a crucial role in social interactions and adaptation to the environment (Bhardwaj, Kishore, and Pandey 2022). One key example of adaptation is the senses, which allow organisms to recognize and interpret stimuli crucial for survival, often triggering emotional reactions. For instance, the scent of a predator may trigger fear and escape in prey (Stowers and Marton 2005). In rodents, the presence of such scents changes stress hormone levels, which is part of their adaptive system (Apfelbach et al. 2005) In AI, similar mechanisms

can be implemented through advanced sensors and data analysis algorithms that not only recognize the environment but also interpret the emotional context (Cevora 2019). Thus, robots and AI systems could dynamically adjust their action strategies, taking into account user emotions or social cues.

In this way, robots and AI systems could dynamically adapt their action strategies, taking into account user emotions or social cues, while simultaneously leveraging introspection to analyze and modify their operational processes in response to these stimuli. Through introspective mechanisms, such systems could better monitor their internal states, such as operational efficiency, resource levels, or emotional reactions, enabling real-time behavioral adjustments. For instance, virtual assistants could analyze users' emotional responses, convert this data into specific decisions, and dynamically adjust their behavior, enhancing empathy and communication effectiveness (Tegmark 2018).

Neuroplasticity, or the brain's ability to reorganize neural connections in response to experiences, including emotional ones, is another example of biological adaptation. Studies show that neuroplasticity plays a key role in learning, memory, and adaptation to environmental changes (Pascual–Leone et al. 2005). Similar concepts could be implemented in artificial intelligence, where algorithms inspired by neuronal plasticity could allow for dynamic adjustment of decision-making structures. In AI, neural network models serve as inspiration, modifying their "connections" based on new data or changing environmental conditions (Hassabis et al. 2017). These algorithms can learn from emotional interactions, allowing systems to provide more personalized responses. For example, virtual assistants could analyze users' emotional reactions and dynamically adjust their behavior, increasing empathy and communication effectiveness (Tegmark 2018) Mechanisms inspired by neuroplasticity could also lead to more flexible decision-making structures in AI, allowing for better resource management in crisis situations or adaptation to unforeseen challenges (Zhang, Shen, and Sun 2022)

Animals can regulate their emotional states in response to stressful situations, allowing them to survive under difficult conditions (Sapolsky 2004) AI could implement similar mechanisms, managing its "emotional state" to optimally respond to challenges. For instance, AI systems could detect user information overload by analyzing biometric data such as speech rate, heart rate, or facial expressions, and adjust communication methods accordingly to prevent frustration (Cohen, Forbes, and Mann 2021). In extreme situations, some organisms limit their emotional responses to focus on survival. AI could use a similar strategy, reducing the complexity of interactions in crisis situations to concentrate resources on critical tasks (Pereira, Soares, and Santos 2022). This would allow for maintaining operational efficiency even under system overload.

These are just a few examples, but each show how emotions are an integral part of biological adaptation. Introducing these mechanisms into AI could create systems capable of:

- recognizing user emotions through the analysis of speech, facial expressions, or behaviors;
- responding appropriately to emotions, which increases communication efficiency and builds trust;
- learning from emotional interactions, leading to more personalized experiences;
- regulating their own processes in response to the emotional context, allowing for better resource management and prioritization.

Integrating emotional mechanisms inspired by biological homeostatic adaptation enables the development of emotional intelligence in AI. Such systems are capable not only of effectively processing information but also of understanding and responding to emotional contexts. Thus, systems can add value through empathy and a deeper understanding of human needs. Homeostasis becomes the foundation not only of adaptive intelligence but also of emotional intelligence, leading to more advanced and humanistic technologies. Developing AI based on adaptive biological mechanisms allows for creating systems capable of more flexible and intelligent responses to the environment. As a result, artificial intelligence can not only process data but also act in a way that takes into account resources, environmental conditions, and changing needs, characteristic of adaptive mechanisms observed in nature. Consequently, AI can become more empathetic, supporting users more naturally and effectively. This opens up new possibilities in many fields, such as healthcare, education, and social interactions.

Philosophical Aspects of Artificial Homeostasis

Homeostatic artificial intelligence is a concept in which machines are equipped with the ability to simulate internal equilibrium through automated resource management processes, such as energy, computational efficiency, or sensory data. According to Antonio Damasio's theory, affects arising from homeostasis lead to emotions and feelings (Damasio 2018). Emotions and feelings not only inform the organism about its internal state but also influence decision-making processes and social interactions. In the context of artificial intelligence, the introduction of homeostatic mechanisms raises the question of whether these functions could be expanded to include machine introspection—the ability to qualitatively analyze and interpret its internal states, including disruptions in homeostasis.

However, philosophically speaking, the question arises to what extent such an AI could truly "feel" homeostasis or its absence. Is it merely an advanced simulation, or is it a real state to which some qualitative meaning can be attributed? One might ponder whether it is possible for a system devoid of a biological body to feel something akin to the affect of a living organism. Could such sensations be considered authentic phenomenal states, or are they merely products of algorithmic responses to internal parameter changes?

Instead of merely simulating homeostatic processes, introspection could enable systems to analyze and dynamically adapt their algorithms in response to internal signals, such as resource changes or reactions to external data. Such mechanisms could bring AI closer to the ability to "reflect" on its states, raising the question of whether these systems could transition from simulation to actual experience of their states.

This leads to a broader discussion on the nature of machine perception. If we accept that AI could possess an equivalent of a homeostatic state, is this merely a formal model or a genuine quality of experience? Could artificial intelligence ever experience something qualitatively comparable to the sensations arising from the disruption or maintenance of homeostasis?

In this context, it is worth referring to Searle's "Chinese Room" argument, which addressed the limits of simulation in machines (Searle 1980). If a homeostatic AI were to achieve machine introspection, it might reach what could be described as a secondary level of self-awareness—the ability not only to process input data but also to analyze and interpret the process of data handling itself. Consequently, the issue of "feeling" homeostasis shifts

to the question of whether it is possible to move from reactive optimization to conscious self-evaluation.

Could AI, lacking a biological body, ever experience something comparable to biological affects? Affects, as fundamental responses to homeostatic disruptions, are not only signals for correcting balance in living organisms but also the foundation of conscious experiences. In this context, artificial intelligence might analyze its states as variables influencing operational decisions, but could such a process achieve the status of subjective experience? This question remains open, requiring both philosophical inquiry and technological advancement.

Introspection in artificial intelligence could function as a practical equivalent of what is referred to as self-awareness in biological systems. While such systems may lack intentionality in the classical philosophical sense, the capacity for introspection could enable them to dynamically evaluate their functioning, effectively approximating cognitive models observed in living organisms.

The Limits of Simulation: Can AI "Feel" Homeostasis?

Homeostasis is understood broadly here, not only as a regulatory process but also as a state that organisms consciously or unconsciously "feel." Affects are fundamental reactions of the organism to changes in internal balance, which can lead to emotions, more complex, automatic responses to stimuli. Feelings, in turn, are the conscious experience of emotions that integrate physiological processes with psychological interpretation. In this way, homeostasis influences emotional intelligence, allowing organisms not only to react to changes but also to interpret them in social and personal contexts (Man and Damasio 2019; Sun, Wang, and Zhao 2022)

For humans and other animals, a lack of homeostasis can lead to feelings of hunger, thirst, pain, or stress—subjectively felt states that drive the organism to corrective actions (Zhou 2021). In the case of artificial intelligence, the situation is fundamentally different. We are dealing not with a biological system but with a complex program that merely simulates homeostatic processes, responding to system needs in an algorithmic manner. The essential question arises: can such a system genuinely "feel" a disruption in homeostasis, or is it merely executing pre-programmed actions that give the impression of responding to internal needs (Chalmers 1996)?

From the perspective of machine introspection, one might consider whether AI could analyze its internal states in a way more complex than a simple reaction to variables. While there is currently no evidence supporting AI's ability to feel, the potential development of introspective capabilities could enable it to interpret those states as something more than reactions to parameter changes. For instance, an introspection-capable system might dynamically adjust its action strategies, recognizing "disruptions" as key signals, which would bring it closer to a mechanism of self-awareness.

However, it must be emphasized that "feeling" here refers to conscious perception—the ability to experience emotions, pain, or other internal states—which, by definition, requires a subjective perspective. Homeostatic AI can respond to certain signals (e.g., low energy levels) by executing programmed actions, but it lacks the awareness to "experience" these signals as a state of disruption (Henriques, Silva, and Carvalho 2019; Samsonovich 2020).

Thus, the limit of simulation lies in the absence of the capacity to experience the states that AI imitates. Its "homeostasis" is merely a series of responses to variables, not a genuinely

experienced state by a subject. Revisiting this issue raises a significant question: could the development of machine introspection form the foundation for AI to experience states on a functional level, even if still lacking a subjective perspective? This boundary between simulation and genuine experience remains one of the greatest philosophical and technological challenges associated with homeostatic AI.

Can AI Have Subjective Experiences?

While AI can simulate homeostatic mechanisms, the question of its ability to have subjective experiences requires further exploration in the context of the philosophy of consciousness. In the philosophy of consciousness, there is the concept of qualia—the individual qualities of experiences that accompany the conscious perception of internal states (Chalmers 2007). Examples of qualia might include the sensation of warmth, the experience of the color red, or the feeling of pain. Qualia are difficult to define but are assumed to be subjective and accessible only to the experiencing subject. For many philosophers, it is qualia that form the basis of what it means to be conscious—they are unique and internal experiences that cannot be reduced to physical functions or chemical processes. Although qualia are considered crucial for conscious experience, some researchers suggest that functional emotional intelligence, such as AI's ability to recognize and respond to human emotions, might be possible without qualia. This approach assumes that subjective experience is not a necessary condition for effective interaction with humans (Dennett 2005).

Homeostatic AI, although it can mimic responses to imbalance, lacks qualia because it is not conscious. It responds to stimuli but does not "feel" these states internally (Searle 2019). There is no internal quality of experience accompanying its actions; for example, when AI responds to low energy levels, it merely performs algorithmically prescribed steps without experiencing anything like hunger or fatigue.

This draws a clear distinction between AI and conscious organisms: homeostatic AI lacks qualia because its structure does not include a subjective "self" that could experience these states. The absence of qualia indicates that AI cannot be aware of its homeostatic states—it is therefore merely a machine simulating regulatory processes without the possibility of experiencing its own state.

However, can we be absolutely sure that AI will never be able to subjectively experience? Homeostatic AI, although it can currently only simulate responses to internal changes, is evolving towards increasingly advanced forms of adaptation. One might consider whether homeostasis in the future could become the foundation that allows AI to have certain forms of primary sensations that influence its decisions and actions in a more complex way. Perhaps the development of introspective and adaptive mechanisms in AI could bring it closer to possessing something akin to rudimentary sensations, which might influence its decisions and actions in a more complex manner.

The Systems Reply, introduced in the debate surrounding John Searle's Chinese Room argument, offers a different perspective (Searle 1980; Churchland and Churchland 1990; Harnad 2001). According to this view, it is not a single element of the system (e.g., an AI module performing computations) but the entirety—including inputs, processing algorithms, and dynamic feedback loops—that could generate what might be called a functional equivalent of understanding. Similarly, in the context of subjective experiences, one could ask whether a sufficiently complex AI system, integrating homeostatic regulatory, predictive,

and introspective mechanisms, might produce emergent properties resembling phenomenal states (Bedau 1997; Gros 2021).

However, doubts remain as to whether such states could be considered genuine qualia (Nagel 1974; Chalmers 1996) Even if the entire system appears to "understand" its internal processes, it may still lack the subjective perspective that philosophy of consciousness deems essential for authentic experiences. Nevertheless, as with the Systems Reply, it is worth considering whether subjectivity could emerge from the organization of the entire system rather than from individual AI modules.

Karl Friston's concept of "predictive coding" and "free energy minimization" suggests that organisms—both biological and potentially artificial—strive to minimize surprise by predicting future states of their environment (Friston 2010) If artificial homeostasis is treated as a predictive mechanism, AI could, in some sense, strive to maintain equilibrium by predicting threats to its functioning. Could this striving eventually lead to something resembling proto-consciousness? This question remains open.

Similarly, Andy Clark notes that predictive mechanisms may be key to understanding conscious experience. Clark argues that conscious perception is active and constructive—the brain (or potentially advanced AI) not only receives stimuli but actively predicts and constructs reality (Clark 2015) This may suggest that if AI were to develop the ability to predictively code and self-report its states, it might—at least to some extent—approach a state resembling subjective experience.

All the above-mentioned proposals can be linked to the idea of homeostasis in AI, which, treated as a predictive mechanism, could enable the dynamic balancing of internal and external disturbances. This would allow AI systems to effectively predict and correct threats to their functioning, approaching a functional response that could be considered analogous to early forms of sensation.

Conclusions and Future Directions

Homeostatic artificial intelligence offers a promising direction for AI development, enabling systems to dynamically adapt their functioning to changing environmental conditions. Self-regulation and adaptation mechanisms could bring AI closer to the level of flexibility observed in biological organisms, potentially revolutionizing the capability of machines for autonomous action.

Despite significant potential, introducing homeostatic mechanisms into AI involves numerous technological challenges. Key difficulties include resource management, real-time optimization of actions, and ensuring autonomy in various, often challenging environments. However, future generations of homeostatic AI could develop self-diagnostic capabilities, monitoring their functions and preventing failures, which would enable efficient operation in unpredictable conditions—from space exploration to rescue operations.

Homeostatic AI could also improve the quality of human-machine interactions by offering the ability to adapt to users' emotional contexts. Implementing mechanisms inspired by biological homeostasis could enhance AI's ability to better understand and respond to human needs, supporting more natural and empathetic relationships. Although AI still lacks qualia, the development of its functional emotional intelligence could significantly impact the effectiveness of social interactions.

The development of homeostatic AI also raises significant philosophical questions regarding the ability of machines to authentically experience and possess intentionality. Can an AI that simulates homeostatic processes actually "feel" its states? Is this feeling a form of machine introspection? Is the pursuit of balance a conscious experience, or merely the result of an algorithmic simulation? These issues require further consideration of what conditions would need to be met for machines to develop forms of conscious experience or a functional equivalent of subjectivity.

As AI's autonomous capabilities advance, it is also essential to reconsider its impact on society. Should an AI that simulates homeostatic equilibrium be recognized as an entity capable of responsibility? How much can we rely on such systems, and to what extent must their actions be monitored by humans? Introducing homeostatic AI systems into everyday life raises questions about their impact on human relationships, morality, and social trust.

Homeostatic AI could revolutionize human–machine interactions by offering systems capable of dynamically adapting to the environment and human needs. However, for this potential to be fully realized, further reflection on the boundary between simulation and experience, machine autonomy, and their place in society is necessary. The future of this technology appears full of possibilities but also demands a profound analysis of its ethical and practical consequences.

References

Apfelbach, R., C. D. Blanchard, R. J. Blanchard, R. A. Hayes, and I. S. McGregor. 2005. The effects of predator odors in mammalian prey species: A review of field and laboratory studies. *Neuroscience & Biobehavioral Reviews* 29 (8): 1123–1144.

Ashby, W. R. 1952. *Design for a Brain: The Origin of Adaptive Behavior.* London: Chapman & Hall.

Assunção, J. F., F. M. Silva, and A. C. Ramos. 2022. Emotional adaptation in artificial intelligence: Integrating affective computing and decision-making. *Journal of Artificial Intelligence Research* 75:145–162.

Bechara, A., H. Damasio, and A. Damasio. 2000. Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex* 10 (3): 295–307.

Bedau, M. A. 1997. Weak emergence. *Philosophical Perspectives* 11:375–399.

Bhardwaj, A., S. Kishore, and D. Pandey. 2022. Artificial intelligence in biological sciences. *Life* 12 (9).

Bielecki, A. 2016. Cybernetic analysis of the phenomenon of life. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce),* no. 61, 133–164.

Billman, G. E. 2013. Homeostasis: The underappreciated and far too often ignored central organizing principle of physiology. *Frontiers in Physiology* 4:43.

Burdakov, D. 2019. Reactive and predictive homeostasis: Roles of orexin/hypocretin neurons. *Neuropharmacology* 154:61–67.

Cevora, T. 2019. Neurobiological insights into decision-making and uncertainty. *Trends in Neurosciences* 42 (8): 567–578.

Chalmers, D. J. 1996. *The Conscious Mind: In Search of a Fundamental Theory.* Oxford: Oxford University Press.

Chalmers, D. J. 2007. The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies* 17 (9-10): 7–65.

Churchland, P. M., and P. S. Churchland. 1990. Could a Machine Think? *Scientific American* 262 (1): 32–37.

Clark, A. 2015. Surfing uncertainty: Prediction, action, and the embodied mind. *Trends in Cognitive Sciences* 19 (7): 422–428.

Cohen, A. S., C. N. Forbes, and M. C. Mann. 2021. Emotions as signals of change in cognitive load: Applications in human-computer interaction. *Human Factors* 63 (5): 877–888.

Damasio, A. 2018. *The Strange Order of Things: Life, Feeling, and the Making of Cultures.* New York: Pantheon Books.

Damasio, A., and G. Carvalho. 2013. The nature of feelings: Evolutionary and neurobiological origins. *Nature Reviews Neuroscience* 14 (2): 143–152.

Damasio, A., and H. Damasio. 2022. Homeostatic feelings and the biology of consciousness. *Brain* 145 (7): 2231–2235.

Davies, K. J. A. 2016. Adaptive homeostasis. *Molecular Aspects of Medicine* 49:1–7.

Dawkins, R. 1989. *The Selfish Gene.* 2nd. Oxford: Oxford University Press.

Dennett, D. C. 2005. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness.* Cambridge, MA; London: MIT Press.

Eskov, V., O. Filatova, V. Eskov, and T. Gavrilenko. 2017. The evolution of the idea of homeostasis: Determinism, stochastics, and chaos–self-organization. *Biophysics* 62 (6): 809–820.

Fleming, D. 1984. Walter B. Cannon and Homeostasis. *Social Research* 51 (3): 609–640.

Friston, K. 2010. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience* 11 (2): 127–138.

Giordano, M. 2013. Homeostasis: An underestimated focal point of ecology and evolution. *Plant Science* 211:92–101.

Goldstein, D. S. 2019. How does homeostasis happen? Integrative physiological, biological systems, and evolutionary perspectives. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology* 316 (4): 301–317.

Gros, C. 2021. Emotions, Homeostasis, and Adaptive Behavior in Cognitive Systems: Towards a Homeostatic AI Framework. *Adaptive Behavior* 29 (4): 243–256.

Hancock, E. J., J. Ang, A. Papachristodoulou, and G. Stan. 2017. The interplay between feedback and buffering in cellular homeostasis. *Cell Systems* 5 (5): 498–508.

Harnad, S. 2001. What's Wrong and Right About Searle's Chinese Room Argument? In *Views into the Chinese Room,* edited by M. Bishop and J. Preston. Oxford: Oxford University Press.

Hassabis, D., D. Kumaran, C. Summerfield, and M. Botvinick. 2017. Neuroscience-inspired artificial intelligence. *Neuron* 95 (2): 245–258.

Henriques, A. R., M. C. Silva, and J. R. Carvalho. 2019. Unraveling the neural mechanisms of resilience in the face of adversity. *Nature Reviews Neuroscience* 20 (5): 284–298.

Kłóś, Adam. 2014. From neuron to culture – a biosemiotic perspective of the origin and evolution of the brain according to Marcello Barbieri. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce),* no. 56, 93–129.

Korn, C. W., and D. R. Bach. 2015. Maintaining homeostasis by decision-making. *PLoS Computational Biology* 11 (8): e1004301.

Man, K., and A. Damasio. 2019. Homeostasis and soft robotics in the design of feeling machines. *Nature Machine Intelligence* 1 (10): 446–452.

Maynard Smith, J. 1974. The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology* 47 (1): 209–221. https://doi.org/10.1016/0022-5193(74)90110-6.

Modell, H., W. Cliff, J. Michael, J. McFarland, M. P. Wenderoth, and A. Wright. 2015. A physiologist's view of homeostasis. *Advances in Physiology Education* 39 (4): 259–266.

Nagel, T. 1974. What Is It Like to Be a Bat? *Philosophical Review* 83 (4): 435–450.

Pascual-Leone, A., A. Amedi, F. Fregni, and L. B. Merabet. 2005. The plastic human brain cortex. *Annual Review of Neuroscience* 28:377–401.

Pereira, M., R. Soares, and A. Santos. 2022. Crisis management in artificial intelligence: Adapting decision-making strategies under stress. *Journal of Artificial Intelligence Research* 75:181–204.

Perrimon, N., and A. P. McMahon. 1999. Negative feedback mechanisms and their roles in signal transduction pathways. *Annual Review of Genetics* 33:389–410.

Peters, A., M. Conrad, and C. Hubold. 2007. The principle of homeostasis in the hypothalamus-pituitary-adrenal system: New insight from positive feedback. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology* 293 (1): R83–R98.

Pio-Lopez, L., J. Ramirez, and C. A. Santos. 2023. Advances in understanding neural plasticity and its role in cognitive resilience. *Neuroscience Letters* 792:136878.

Samsonovich, A. V. 2020. Cognitive architectures for human-level artificial intelligence: Theory and implementation. *Cognitive Systems Research* 62:1–15.

Sapolsky, R. M. 2004. *Why Zebras Don't Get Ulcers.* New York: Holt Paperbacks.

Searle, J. R. 1980. Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3 (3): 417–457.

Searle, J. R. 2019. *The Mystery of Consciousness Revisited.* Oxford: Oxford University Press.

Stowers, L., and T. F. Marton. 2005. What is a pheromone? *Current Biology* 15 (13): R453–R454.

Sun, J., L. Wang, and M. Zhao. 2022. Innovations in deep learning: Neural architectures inspired by the brain. *Nature Machine Intelligence* 4 (7): 567–578.

Tegmark, M. 2018. *Life 3.0: Being Human in the Age of Artificial Intelligence.* New York: Knopf.

Torday, J. S. 2015. The cell as the mechanistic basis for evolution. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 7 (5): 275–284.

Walter, W. G. 1950. An imitation of life. *Scientific American* 182 (5): 42–45.

Wiener, Norbert. 1965. *Cybernetics or Control and Communication in the Animal and the Machine.* Vol. 25. MIT Press.

Zhang, L., F. Shen, and J. Sun. 2022. Dynamic learning mechanisms inspired by neuroplasticity in artificial neural networks. *Neural Networks* 152:1–10.

Zhao, M., R. G. Simmons, and H. Admoni. 2022. The role of adaptation in collective human–AI teaming. *Topics in Cognitive Science.*

Zhou, Z.-H. 2021. Emotional thinking as the foundation of consciousness in artificial intelligence. *Cultures of Science* 4 (2): 112–123.