The Cybernetic Revolution in Science Fiction: Theoretical Prefiguration and Ethical Transcendence in the Planetary Imagination of Solaris

Ruien Zhang*
East China Normal University, Shanghai 200062, China
*Corresponding email: 18516123813@163.com

Abstract

Stanisław Lem's *Solaris* stands as a paradigmatic work embodying the profound interplay between science fiction literature and cybernetic thought. By constructing the grand literary thought experiment of the planet Solaris, the novel rigorously demonstrates classical cybernetic models. It also profoundly exposes the limitation of these models - positioning the observer outside the system - through its paradoxes in communication and teleology. Lem's literary imagination exhibits remarkable foresight. On the one hand, the issue of "observer intervention" directly foreshadows the theoretical revolution from first order to second-order cybernetics. On the other hand, the extreme complexity of the planet itself resonates deeply with later new paradigms in systems science, including autopoietic theory and dissipative structure theory. Ultimately, the core value of the novel lies in its transcendent ethical contemplation. When attempts at "control" prove utterly futile, the narrative shifts to the fundamental question of how to coexist with an absolute "Other". By returning to the dynamic adaptive wisdom inherent in the original metaphor of cybernetics, the "helmsman", Lem's work criticizes the technical rationality that reduces cybernetics to a mere tool of domination. In doing so, it elevates and transcends cybernetic thought from a scientific model to a philosophical and ethical discourse.

Keywords

Stanisław Lem, Solaris, Cybernetics, The Other

Introduction

Cybernetics, the science that studies control and communication within and between living organisms, machines and organizations, emerged and evolved in the 1950s. It exerted a profound influence on numerous scholars of the contemporary era and subsequent generations. It has not only established the laws governing the transformation of regulatory information but also built a bridge connecting traditional humanistic disciplines and thermodynamics. It has also been theoretically expanded and rediscovered as a novel conceptual framework and tool model across diverse academic fields. In the ongoing process of mutual interaction between cybernetic thought and literature, Stanisław Lem. The renowned science fiction writer and member of the Polish Cybernetics Association played a pivotal catalytic role. As Joanna Zylinska pointed out, "It is interesting to regard Lem as a conscious and critical practitioner of cybernetics as a theoretical framework" [1]. In fact, he possessed the core competencies of a cybernetics expert. In *Summa Technologiae*, he repeatedly employed core cybernetic concepts to analyze social, technological and evolutionary issues.

At the same time, however, he maintained a fundamental skepticism and reflection toward cybernetics, and attempted to "introduce moral issues into cybernetic thought" to conduct ethical contemplation. Notably, the profound connection between Lem and cybernetics is reflected not only in his theoretical works. It is also embedded more deeply in his literary creations, with Solaris, published in 1961, standing out as the most paradigmatic example. While most domestic scholars still focus their research on Solaris' unique narratological logic and its challenge and subversion of anthropocentrism. Some scholars have identified the hidden links between Solaris and

cybernetics. They have also conducted detailed analyses and analogies between the planet and the human brain by drawing on the core cybernetic concept of the "black box". Thus, it attracts academic attention to the close and intricate interweaving between cybernetic science and literature [2].

In this sense, science fiction literature not only re-presents and responds to scientific theories, but also envisions and experiments with future science. It shoulders the mission of ethical reflection and moral reconstruction at the same time. Therefore, cybernetics is continuously expanding into the humanities in the contemporary era. It has become an urgent academic endeavor to thoroughly examine the interaction and symbiosis between science and literature at the level of deep narrative and ideology. This research is carried out from an interdisciplinary perspective. As one of the most classic and compelling texts in this regard, *Solaris* will provide substantial support for such research.

Correspondence and paradox: The dual narration of cybernetics in *Solaris*

Stanisław Lem possessed profound insight into and critical reflection on cybernetic models. Such thinking is distinctly manifested in Solaris. It is a perfect control system capable of self-regulation, operation and stable existence. In terms of its own operation, survival and stability, Solaris exhibits the fundamental characteristics of a control system and conforms to the basic model of classical cybernetics. Most notably, it features a negative feedback regulation mechanism. The purpose of negative feedback is to maintain the system's inherent stability. Subject to gravitational forces, Solaris should have deviated from its orbit and "been swallowed by the blazing star". Yet "in reality, its orbit of revolution is remarkably stable", as if it "can actively influence the planet's orbital movement" to achieve self-correction after deviation [3]. In other words, the oceanic substances affected by gravity can flexibly change and modify the gravitational pull of the host star in some way. They stabilize the planet's orbit in this manner. This fulfills the goal of self-preservation [4].

This negative feedback mechanism perfectly aligns with the essential attribute of control systems - "to determine and adjust their own movements based on certain changes in the surrounding environment". It lays the foundation for the planet's steady operation [5]. A control system is not an ordinary machine that repeatedly performs monotonous actions. It must possess a certain degree of flexibility and adaptability. Whether through the mimoid, which instantly replicates inanimate objects such as puppets and tools in the environment. Or through the endlessly varying symmetrical and asymmetrical cones that reorganize according to spatiotemporal parameters. Solaris demonstrates extraordinary transformability. It can make flexible behavioral adjustments in response to environmental conditions, and achieve a certain level of learning and evolution amid human research and stimulation. In addition, communication and information transformation are fundamental characteristics of control systems, and the information flow within Solaris is exceptionally complex. Complex information emitted by human observers, such as radiation detection signals and brainwaves, is received, stored, and processed by Solaris. Which ultimately converts them into outputs such as "visitors", symmetrical cones, or gravitational variations.

As Snowdon put it, "it extracts the 'spark of memory' from the depths of our minds and records it in exhaustive detail". The acquisition, utilization, retention and transmission of information take place continuously on Solaris. Although the intricate process involving Snowdon eludes human observation and research, it precisely highlights Solaris' consistency with the cybernetic "black box" model. Incapable of being algorithmically programmed, with its internal operational mechanisms impossible to fathom or observe, and its future states unpredictable. Solaris can only be studied "by relying on input and output information" and further "inferring the complex connections that may exist within it". Yet a complete grasp of its structure remains unattainable. Thus, the enigmatic Solaris is on the one hand an unknowable ocean. On the other hand, it conforms to the basic model of cybernetics, presenting itself as a standard control system.

Nevertheless, as Darko Suvin defined science fiction: "The necessary and sufficient condition for a literary genre to be science fiction is the presence and interaction of estrangement and cognition" [6]. The purpose of science fiction is not mere popularization of

scientific knowledge, nor is it pure fantasy. Instead, it constructs a defamiliarized setting through sound scientific logic. It thereby reflects on those taken-for-granted, unquestioned issues in the real world. Although Solaris' negative feedback homeostasis, environmental adaptability, and black-box information processing mode perfectly align with the characteristics of a control system. It still embodies contradictions that transcend the classical cybernetic model.

This is distinctly manifested in the "communication" between Solaris and humans. The lack of improvement in the F-products indicates the absence of a feedback correction loop. Sartorius' analysis accurately identifies the anomaly of Solaris. The human elimination of "F-products" fails to prompt the system to adjust its behavior in response to such "errors". This phenomenon of ignoring the repeated destruction of F-products and refusing behavioral optimization seems paradoxical given the planet's precise orbital regulation and vivid imitation of tools. It leaves researchers uncertain whether Solaris can learn from past experiences and memories.

Meanwhile, communication between humans and Solaris suffers from channel misalignment. This is due to the incommensurability between scientific signals such as electromagnetic waves and detectors, and non-symbolic entities such as F-products and mimoids. Even the only observable F-products are regarded as indeterminate subjects due to their contradictory state of being both human-like and non-human, tangible and intangible, independent and dependent simultaneously. In this sense, Lem indeed cut off all clues leading to the incarnational dimension of the creation, so as to prevent contact with the Solaris ocean from following the pattern of human interaction. He also rendered ineffective any conventional communication and adaptation between systems [7]. More fundamentally, according to Norbert Wiener's definition, a control system operates toward a specific goal and can "adopt specific functions based on past experience".

However, Solaris merely maintains physical homeostasis and appears to have no purpose whatsoever. The precise ocean would never conduct such crude experiments. Therefore, the F-products only project aspects of humans themselves and lack negative feedback correction. The asymmetrical cones generate irregularly and repeatedly collapsed moments after formation. These phenomena can only be meaningless self-amusements of the system. These phenomena corroborate Kelvin's metaphor of Solaris as an "infant": It saves nothing, serves nothing; it merely exists, existing in an unconscious, a priori, absolute and self-evident manner, much like a baby [8]. This aimless and non-functional Solaris transcends the applicable boundaries of the classical cybernetic model, and renders the black-box model absolutely closed. Observers cannot only fail to analyze its structure through input and output information, but also struggle to judge the "intentions and efficacy" of Solaris as a control system.

Scientific prophecy: The echo between planetary imagination and the development of cybernetics

The contradictory state of Solaris, a body that both conforms to and deviates from a control system, stems from two major causes. The first is the intervention of observers. Unlike the assumption of classical cybernetic models, the experts on the spaceship are neither completely independent of the system nor superior to it. On the contrary, the observers' memories, subconsciousness and even unconsciousness serve as the raw material of Solaris. The "visitors" are formed by the reverse assimilation of the observers' behaviors. They prove through extreme examples that observers cannot achieve independence, nor can they prevent themselves from influencing the system [9].

Therefore, observers become part of the system or even its variables. The distorted negative feedback, ineffective communication and adaptation, and the aimless, functionless black box all exist as a result. These factors render the control model inapplicable. Notably, the questions raised in *Solaris* (1961) received a response in the keynote speech at the 1979 Fall Conference of the American Society for Cybernetics. Heinz von Foerster put forward the concept of the "second-order deficiency" encapsulated in the phrase "we cannot see what we do not see". He also proposed a corresponding "second-order remedy": We must observe our own observing behavior, and ultimately explain the explanations we put forward [10]. In other words, the focus must shift from observing external things to observing the act of "observing" itself. This

shift enables "cyberneticians to enter their own (cybernetic) domain" and thereby identify and interpret their own activities. The cognitive shift in the role of observers, coupled with the evolution of related theoretical models, has led to a clear division of cybernetics into two phases. The earlier phase is "first-order cybernetics, the cybernetics of observed systems", while the later phase is "second-order cybernetics, the cybernetics of observing systems" [11]. If, as the second-order cybernetics school advocates, the development from first to second-order cybernetics constitutes a "revolution" in cybernetics. It is comparable to Einstein's theory of relativity vis-à-vis Newtonian classical mechanics [12]. Then Lem's *Solaris* should be regarded as pivotal work in this revolution.

Beyond the factor of observers themselves, the complexity of Solaris is another crucial reason for the failure of classical cybernetic models. On one hand, the ocean exhibits a constant state of self-reference and recursion. Inside the symmetrical cones, "propagation occurs everywhere, with new embryos emerging like avalanches". The colloidal substances themselves undergo "continuous or periodic dissolution and reconstruction, creation and destruction, production and consumption", engaging in a form of self-organization or "autopoiesis" [13]. Not only matter but also energy is in a perpetual cycle of generation and consumption. The fluctuating ocean constantly undergoes cycles of order, disorder and renewed order. On the other hand, the ocean, with its vast array of material components and ultra-high information transmission density, maintains a state that is both chaotic and precise, deterministic and random. As Kelvin concludes in the end, "There is an indescribable and stark contrast between the agility and curiosity of that living being, and the boundless, far-reaching expanse of the ocean".

Whether it is the "visitors" that appear randomly yet perfectly mirror the observers' memories. Or the symmetrical cones that form perfect symmetry without discernible rules or logic, or the mimoids that accurately replicate tools and objects but fail to imitate humans. Solaris embodies a profound contradiction between microscopic agility and precision, and macroscopic chaos and vastness. This contradiction is difficult to comprehend and cannot be fully explained by traditional methods such as statistics. However, the complexity and dilemmas posed in *Solaris* have both been analyzed and interpreted in cybernetics and systems theory since the 1970s. The autopoietic theory proposed by Chilean biologists and cognitive scientists Humberto Maturana and Francisco Varela presents a model of self-organizing systems. These systems feature self-reference and recursion capable of achieving self-production and thus self-maintenance. This model bears a certain isomorphism with Solaris. Meanwhile, the core concept of "living ordered structures" in the dissipative structure theory founded by Ilya Prigogine is clearly manifested in the fluctuations of matter and energy on Solaris. In addition, the Complex Adaptive Systems (CAS) theory proposed by John H. Holland in the 1990s is characterized by "the unity of randomness and determinism", which also shows a distinct echo with Solaris [14,15].

Therefore, we might argue that through *Solaris*, Lem conducted a grand literary thought experiment. He not only applied the cybernetic models of his era with precision. He also, through profound insight, anticipated the fundamental challenges the theory would face in future development: The observer paradox and the problem of system complexity. The autonomous control system of Solaris he constructed in the novel provides a highly prescient literary rehearsal and philosophical inquiry for the subsequent development of cybernetic theory. In this sense, Lem's science fiction creations transcend mere literary responses and greatly enrich and expand the ideological territory and application context of cybernetics.

The return of ethics: Wisdom from the failure of control to coexistence with the other

Stanisław Lem's literary thought experiment through Solaris not only anticipated the impending revolution within cybernetic theory, but also inevitably directed cybernetic inquiries toward the ethical dimension. Its core lies in the repeatedly addressed concept of "observer intervention". When the observer exists as an inseparable variable within the system, they introduce not only deviations in scientific cognition but also projections of their own desires and purposes. The underlying drive behind human attempts to name, classify, model and communicate with Solaris is a

deep-seated will to control, an urge to incorporate the planet into a manageable cognitive framework. Yet Solaris, with its absolute autonomy and incommensurability, thoroughly frustrates this ambition. It compels human observers to confront a more fundamental question: when control is no longer possible, how should we coexist with an absolute Other? This very inquiry allows *Solaris* to achieve an ethical return to the essence of cybernetics.

Cybernetics is the science of control and communication in animals and machines. Its original purpose was to discover and understand the regulatory capabilities of systems, and to reflect on and analyze communication, feedback, adaptation and system stability. However, due to Norbert Wiener's involvement in the renowned "anti-aircraft gun predictor" project during World War II, the escalating Cold War, and the demand for military equipment and weapons. Cybernetics gradually moved beyond a universal theoretical framework or philosophical perspective. It became a control tool with high military application value, oriented toward human domination of systems [16]. Human research on Solaris is largely built on the core idea of mastery and control. Whether naming the planet's wonders with linguistic symbols such as "slender tendrils" and "tidal surges". Or using metaphors and models to expand boundaries for classification or hypothesizing about potential communication between Solaris and humans. All these signs indicate that human thinking still circles within the confines of Earth and human-centric concepts.

The fundamental intention behind thinking, choosing and acting through these conceptual frameworks is to integrate Solaris into the human cognitive system. Researchers strive to decipher its operational laws and logic through observation and study, aiming to regulate, utilize and even manipulate the planet. But Solaris, as a self-organizing, self-generating autonomous control system, shatters these unrealistic human fantasies [17]. In other words, Solaris has no need to cater to human cognition on human terms. The very existence of the system is its own purpose, and its behaviors are emergences of that existence. Thus, Snowdon's assertion finds validation: We will probably remain ignorant about Solaris, but perhaps we can understand ourselves more thoroughly. Any statement made by the observer is fundamentally a statement about the observer themselves. When our attempt to understand the Other ultimately leads us to recognize our own cognitive limitations, emotional needs and desire for control. This failed communication ushers in a profound process of self-reflection.

Against this backdrop, the perspective of second-order cybernetics offers not only theoretical evolution for cybernetics but also a tool for self-reflection. As "observing systems" endowed with recursive and self-referential capabilities, humans find their thoughts, language and actions exerting cyclic effects on themselves. These shape the realm of their future possibilities. This is analogous to how the Visitors, manifestations of the observers' own thoughts, language and actions, lead the observers to make different choices. This insight directly gives rise to a new ethical consideration. Drawing on this, cybernetician Heinz von Foerster derived a key ethical principle: always acting to increase the number of choices available. For a self-referential system [17,18].

Therefore, actions that narrow the scope of choices, such as stubbornly attempting to conquer or eliminate the Other, are logically self-destructive. They will ultimately erode the foundation of the system's own existence. Conversely, expanding the range of possibilities forms the basis for the long-term survival and freedom of all systems. In this sense, when facing the undecidable question of how we should coexist with an incomprehensible and uncontrollable Other, we attain true freedom. We can choose who we want to be: citizens of an independent universe who fail in their attempt to discover the world's laws, rules and customs. Or conspirators participating in the invention of those very laws, rules and customs. At this point, ethics within second-order cybernetics transforms from external moral norms into an inherent necessity for actors, one that bears on the survival and prosperity of the system. To expand our own long-term existence and space for choice, we must simultaneously expand the living space of the Others with whom we interact. Coexisting with the Other and acknowledging its autonomy is therefore the most rational strategy we, as conscious beings, must adopt to uphold the supreme value of our capacity for choice.

Thus, Solaris, a being that saves nothing and serves

nothing, becomes an equal existence to Earth and humanity, sharing the universe as witnesses to each other's being. True coexistence does not require shared goals or mutual understanding. Kelvin's quiet gesture of reaching out his hand to the ocean before his departure, and the ocean's act of perfectly replicating the shape of his hand before losing interest and receding gently. This stands as an attempt at such coexistence. Similarly, the future of cybernetics may not lie merely in designing or controlling more sophisticated equipment and weapons. Instead, it may lie in learning to maintain reverence in the face of unknowability, and to embrace disinterested reconciliation and acceptance. As Norbert Wiener pointed out in *The Human Use of Human Beings*, progress brings not only new possibilities for the future but also new constraints. Progress symbolized by enhanced communication and technological iteration is also a true reflection of increased human domination over nature [19].

Unrestrained exploitation has thoroughly transformed our environment to such an extent that we must transform ourselves to survive in this new setting. In this way, Lem's narrative leads us back to the original metaphor of the cybernetics' etymological root, the Greek word *kubernetes* meaning helmsman. The wisdom of a helmsman does not lie in conquering or controlling the ocean. It resides in keenly sensing the dynamic shifts of wind and waves and continuously adjusting the rudder to achieve harmonious coexistence with the ocean while maintaining the vessel's course. Its core is dynamic adaptation, not absolute domination.

In this sense, there is a distinct resonance between the ocean faced by the helmsman and Solaris faced by the observer. Kelvin pilots his spacecraft into Solaris, merging with the liquid, eyeless ocean amid growing bewilderment. This form of dignified, dynamic coexistence rooted in reverence not only reveals a deeper, more ethically meaningful potential of cybernetic thought. It also delivers a profound critique of the technical rationality that reduces cybernetics to a mere tool of domination.

Conclusion

Stanisław Lem's science fiction creation is embedded in the historical context of moral and ethical reconstruction in the 20th century. Confronted with the reality that all ethical types inherited from tradition were rapidly losing their validity, science fiction has constantly emerged as a critical field for responding to the predicaments of modernity. As Lem pointed out in a 1994 discussion with Świeszkowski, humanity after the Renaissance has undergone the impacts of great transformations. These include the Enlightenment, the Industrial Revolution and socialism. The rise of extreme nationalism and extreme religious fundamentalism has rendered nihilism a prevailing mood of the era. Therefore, Lem distanced himself from the Enlightenment idea that humanity can transcend tribalism and build a better future on a global scale. Instead, he turned his attention to the pervasive and seemingly unstoppable impulse toward conflict and aggression inherent in human nature. It was precisely this vigilance toward the depths of human nature that prompted Lem to enter a dialogue with cybernetics, a core theory of modernity. Cybernetics is not only concerned with the regulation of technical systems but also deeply linked to the logic of social control and the organization of human behavior.

In this sense, *Solaris* does not stop at a literal literary translation of cybernetics. On the one hand, it vividly demonstrates the path of in-depth interaction between literature and science. Literature is not content with surface-level interpretations of theories. Instead, through narrative experiments. It places scientific propositions at the limits of their cognitive boundaries to conduct extreme tests. In doing so, it reveals the inadequacies of these theories when addressing dimensions such as complexity, consciousness and ethics, and drives them to elevate into deeper philosophical questions. On the other hand, through such narrative experiments, science fiction has also greatly participated in and shaped the innovation of modern ethical concepts.

It forces the focus of ethics to shift from recognition based on identity to responsibility in the face of alterity. This shift transcends the simple critique of anthropocentrism and reshapes the relationship between self and Other at the ontological level. More profoundly, these thought experiments carried out in the context of the 20th century. They have shown remarkable foresight in the present era when cybernetic thought is experiencing a full-scale revival amid the development of artificial intelligence. As we continue to encounter

similar communication paradoxes, questions of alterity and limits of control on the path of AI advancement, Lem seems to have become a visitor emerging from the depths of our minds. He compels people to continuously reflect on their position in technological civilization and the ethical boundaries of their actions. *Solaris* thus proves that outstanding science fiction literature is itself a dynamic form of thought capable of shaping the scientific discourse of the future.

Funding

This work was not supported by any funds.

Acknowledgements

The author would like to show sincere thanks to those techniques who have contributed to this research.

Conflicts of Interest

The author declares no conflict of interest.

References

[1] Konior, B. (2023) Artificial Intelligence in Stanisław Lem's *Summa Technologiae*. *Imagining AI: How the World Sees Intelligent Machines*, 89.

[2] Jaźniewicz, W. (2021) Stanisław Lem's *Solaris*: interpretations in the Russian-speaking world. *The Routledge World Companion to Polish Literature*, 326-335.

[3] Middleton-Kaplan, R. (2020) Refractions of holocaust memory in Stanisław Lem's science fiction. *The Palgrave Handbook of Holocaust Literature and Culture*, 287-304.

[4] Subbotsky, E. (2025) Phenomenology of living consciousness in Stanislaw Lem's novel *Solaris*. *The "Natural Light" of Consciousness: Living Consciousness as a Means and Subject of Psychological Research*, 97-114.

[5] Stichweh, R. (2016) Automaten: Über Norbert Wiener, cybernetics or control and communication in the animal and the machine (1948). *Schlüsselwerke der Systemtheorie*, 19-26.

[6] Terrone, E. (2021) Science fiction as a genre. *The Journal of Aesthetics and Art Criticism*, 79(1), 16-29.

[7] Majewski, P. (2023) Stanisław Lem: Traumatized prophet of posthumanism. *The Polish Review*, 68(2), 142-157.

[8] Grabarczyk, P. (2020) What does the silent planet tell us? The analysis of selected philosophical themes found in Stanisław Lem's Solaris. *Acta Universitatis Lodziensis. Folia Litteraria Polonica*, 59(4), 69-80.

[9] Richards, L. (2018) New directions in second-order cybernetics. *New Horizons for Second-order Cybernetics*, 359-372.

[10] Nowzari, C., Garcia, E., Cortés, J. (2019) Event-triggered communication and control of networked systems for multi-agent consensus. *Automatica*, 105, 1-27.

[11] Sayre, K. M. (2023) Cybernetics. *Philosophy of Science, Logic and Mathematics in the 20th Century*, 292-314.

[12] Chepin, E. (2021) Robotics: From first-order cybernetics to third-order cybernetics. *Procedia Computer Science*, 190, 130-136.

[13] Žukauskaitė, A. (2020) Gaia theory: between autopoiesis and sympoiesis. *Problemos*, (98), 141-153.

[14] Huang, F., Lin, Y., Zhao, R., Qin, X., Chen, Q., Lin, J. (2019) Dissipation theory-based ecological protection and restoration scheme construction for reclamation projects and adjacent marine ecosystems. *International Journal of Environmental Research and Public Health*, 16(21), 4303.

[15] Kunjir, A. (2024) Exploring the applications of complex adaptive systems in the real world: a review. *Artificial Intelligence, Machine Learning and User Interface Design*, 136-160.

[16] Yablonskaya, L. V. (2023) Innovative cybernetic approach in learning process. *SHS Web of Conferences*, 164, 00108.

[17] Potocka-Woźniak, A. (2022) Scenariusze komunikacji z Obcym na podstawie wybranych dzieł Stanisława Lema. *Konteksty Kultury*, 19(3), 407-420.

[18] Lepskiy, V. (2018) Evolution of cybernetics: philosophical and methodological analysis. *Kybernetes*, 47(2), 249-261.

[19] Hancock, P. A. (2019) The humane use of human beings? *Applied Ergonomics*, 79, 91-97.