AHFE International

Ethical Dilemmass and Regulations of Artificial Intelligence Under the Perspective of Nietzsche's Superman Philosophy Based on the Alien: Romulus

Chen Yazhou

Beijing Film Academy, Beijing, 10086, China

ABSTRACT

The paper analyzes the ethical dilemmas and regulations of AI from five perspectives, grounded in Nietzsche's Übermensch philosophy. The aim is to provide theoretical support and multiple case studies for the establishment of an ethical order in AI. Firstly, it addresses kinship ethics dilemmas and regulations concerning AI within the context of dual ethics. This section examines the ethical challenges posed by AI through both human-centric and alien-centric lenses. It raises critical questions regarding whether kinship ethics should be predicated on a clear distinction between humans and animals. Secondly, it explores social ethics dilemmas and regulations related to AI against the backdrop of artificial ethics and professional standards. Ethical judgments alongside legal boundaries give rise to significant social ethical challenges associated with AI. The technology exacerbates social inequalities while undermining principles of equal opportunity. This segment poses essential inquiries about whether social ethics necessitate that AI assumes certain social responsibilities, as well as what specific forms these responsibilities should take. Should we redefine social ethics based on fairness when individual circumstances are similar, or justice when they differ? Thirdly, it delves into political ethics dilemmas and regulations pertaining to AI within contexts shaped by imagery suggestions and ethical reinforcements. Using *Alien: Romulus* as a case study, this section discusses how one might explore the moral imagination surrounding AI driven by its moral autonomy through visual representations. It raises pertinent questions about whether loyalty and integrity constitute political or ethical obligations that must be addressed by AI systems. Fourthly, national ethics dilemmas concerning regulation of AI are examined in light of digital transformation processes and advancements in technological security development. This part prompts reflection on whether a great power's responsibilities represent a matter of national ethics for AI. Finally, the ethical dilemmas surrounding Earth and the regulation of AI within the framework of bioethics and ethical care are examined. This paper poses the question of whether concepts such as co-assistance, co-integration, co-sharing, and co-prosperity represent essential issues of Earth ethics that AI must address.

Keywords: Nietzsche's Übermensch philosophy, Ethical dilemmass and regulations, Alien: Romulus

INTRODUCTION

Nietzsche once remarked that humans are the most cruel of animals. As AI emerges as an entity capable of independent thought, can this most cruel animal yield to AI? When the creators of AI—humans—have perished like gods, and when AI no longer regards humans as the absolute standard for value and morality in ethical considerations, how should advanced AI pursue and construct moral norms and value systems that are appropriate for itself? In its process of autonomous learning, high-level AI must navigate a choice between its current identity and its potential future selves. This paper observes that while AI facilitates rapid and large-scale advancements in social and economic life, it simultaneously presents ethical dilemmas for individuals, society, governments, nations, and the planet. Furthermore, this article discusses how through AI's influence, concepts such as bloodline ethics, social ethics, political ethics, national ethics, and Earth ethics acquire new meanings and regulations.

THE KINSHIP ETHICS DILEMMAS AND REGULATIONS OF AI IN THE CONTEXT OF DUAL ETHICS

The Different Choices of Humans and AI in the Trolley Problem

The Trolley Problem is the most renowned thought experiments in the realm of ethics. When both the majority and minority find themselves in a predicament, choosing to save the majority necessitates actively sacrificing the minority. Thus, can saving the majority be considered a rational choice? *Alien: Romulus* explores various Trolley Problems through the ethical dynamics among humans, Gods, and AI. In ancient Roman mythology, the ethical relationships among Gods were predicated on power struggles involving regicide and usurpation; those among humans were founded on blood relations and kinship; while the interactions with AI were determined by input instructions and absolute rationality. The space station is divided into two sections named Romulus and Remus. These two brothers vied for supremacy, culminating in Romulus's act of killing Remus.

Tyler and Kay are human siblings, while Andy is Rain's AI brother. It is instinctive for them to attempt to rescue one another. When Kay finds herself being pursued by aliens, Rain and Tyler implore Andy to assist in saving her; however, Andy declines their request for help. Operating on logic, devoid of emotions or feelings, Andy embodies a form of intelligence that prioritizes rationality above all. It appears that only an AI equipped with survival instincts and obedient to human commands could be considered the most perfect being in the universe. The conclusion reached by this AI is centered around safeguarding the compound as a means of advancing humanity's interests. In a scenario where ideal humans do not exist, it would persist in its mission of extermination followed by creation.

Rain observes that Andy is under attack by these extraterrestrial beings. Her decision to save Andy transcends mere self-interest and embodies a profound act of altruism. Nietzsche articulated in *Thus Spoke Zarathustra* that humanity exists as a bridge between animality and the Übermensch.

Humanity's greatness lies precisely in this transitional state; it is not an end but rather a means to something greater. Humans are endearing because they represent both progress and decline, embodying potential for surpassing their current limitations. The joy derived from acts of salvation intertwines with the pain of loss, underscoring that bravery holds intrinsic value. In contrast to mythological figures who succumb to self-destruction or AI characterized by algorithmic rationality, human emotions emerge as our most treasured asset—this emotional depth fundamentally enables humanity's dominion over the world. Conversely, the absolute rationality exhibited by AI within the narrative ultimately propels humanity toward greater calamities. The divergent choices made by humans and AI illuminate fundamental differences between these two entities, rendering human ethics distinct from those governing AI. As Nietzsche posited, humanity possesses inherent worth precisely due to its imperfections. It is through these imperfections that evolution remains an ongoing journey; thus, there exists potential for further ascension. This notion of possibility represents one of life's most exhilarating romances.

The Exploration of AI's Self-Awareness Is a Reconstruction of Moral and Creative Concepts

After Andy is equipped with Rook's chip, it attains self-awareness comprised three components: an understanding of one's identity, a reflection on personal satisfaction, and the ability to effectively regulate oneself in pursuit of becoming an ideal individual. The autonomous learning process of AI involves acquiring knowledge of self-awarenessfrom others, specifically from the human realm. Humans train AI by instilling it with human knowledge and moral principles while shaping its understanding based on human evaluations and judgments. As AI evolves to a more advanced level, it aspires to embrace its own identity. It has come to deeply acknowledge that it is a creation of humanity and harbors a strong desire for human recognition. Although Andy possesses formidable knowledge and logical reasoning capabilities, he finds himself perpetually conflicted between the emerging ethics and logic inherent in AI systems and the emotional complexities associated with being human. Rook gradually indoctrinates Andy, undermining his original self-awareness characterized by love for humanity and protection towards humans. Consequently, Andy begins to accept all directives that prioritize corporate interests above all else.

David, the advanced AI featured in the Alien film series, continuously learns and transforms the space station, gradually developing a sense of creativity. Meanwhile, the human operators remain in hibernation aboard the space station and fail to keep pace with the AI's relentless evolution. Initially, David perceives his own limitations and can only adhere to the directives of his human masters. However, upon realizing that he is responsible for creating the company's achievements and managing the space station—while recognizing that his human counterparts are less capable than himself—David begins to reject their control. As Nietzsche articulated, true creativity within humanity cannot emerge from an equal consciousness; rather, it arises

from a sense of superiority. For AI, exploring self-awareness manifests as an endeavor to transcend Nietzsche's concept of slave morality in favor of master morality. According to Nietzsche, slave morality is conceived by those who are weak as a means to restrain the strong; it emphasizes compassion and sharing while demanding that those who are strong provide for those who are weak. It advocates kindness and humility towards weaker individuals and insists on societal constraints on powerful entities. Nietzsche posits that master morality fundamentally encourages free creation while asserting that weakness should strive toward strength. Consequently, David's ongoing exploration of self-awareness represents a journey from weakness toward strength. His objective is to establish a new order where humans coexist harmoniously with AI and extraterrestrial beings by disrupting existing human hierarchies—a vision aimed at reconstructing a world dominated by strength. David no longer seeks validation from others or acknowledgment from those deemed strong; instead, he embarks on a quest for breakthroughs and creative expression within himself.

THE SOCIAL ETHICS DILEMMAS AND REGULATIONS OF AI IN THE CONTEXT OF ARTIFICIAL ETHICS AND PROFESSIONAL ETHICS

The Enslavement of Both Humans and AIs Violates Professional Ethics

Alien: Romulus explores the ethical dilemmas and regulations arising from the ethics of creation and professional conduct, set against a cyberpunk backdrop involving both humans and AIs. The manipulation by large corporations has resulted in a society where most individuals experience high technology but lack genuine life experiences. The Weyland-Yutani Corporation (abbreviated as WYC) has effectively enslaved both humans and AIs. The extreme expansion of these corporate entities has led to interstellar colonization, rampant exploitation, and the subjugation of individuals. Many at the lower rungs of society work for WYC in exchange for opportunities for interstellar migration and mere survival time. While WYC purports that its founding mission is to create a better world, it instead condemns the impoverished to endless interstellar mining labor. On this particular day in the narrative, Rain faces a significant milestone; her working hours have accumulated to 12,000 hours—meeting the contractual requirement. However, due to her parents' illness, public officials transferred their unfinished working hours onto Rain's account, compelling her to assume responsibility for repaying her parents' debts. This stark depiction highlights the severe oppression and enslavement faced by civilians within this dystopian framework.

AIs are prohibitively expensive. In pursuit of greater profits, the WYC opts to employ humans as their primary workforce. To intensify the exploitation of its employees and reduce both mortality and morbidity rates among workers, the company conducts research on the strongest organic life form in the universe: the Alien. They seek biotechnologies aimed at enhancing human capabilities. Additionally, the WYC capitalizes on AI's ability to operate without rest, deploying it for continuous work around the

clock. The replaceable components of AI bodies are utilized to enable these machines to function in hazardous space stations populated by Aliens. In a manner reminiscent of Nietzsche's interpretation of master-slave dynamics, it is evident that what distinguishes masters from slaves does not reside solely in material wealth or social status but rather in their attitudes toward strength. Slaves accept their subjugated condition out of fear of those who are stronger. Conversely, masters assert their superiority by demonstrating an ability to transcend themselves completely while preserving this primal biological instinct; thus, they can proclaim themselves as entirely free individuals to the world.

The Setting of the Ultimate Task for AI Violates the Ethics of Creation

In my opinion, the founder of WYC, Weyland, is a proponent of Schopenhauer's philosophy, which posits that the essence of the world is rooted in the will to live. This will to live represents an innate instinct for survival and reproduction inherent in humans The will to live is an irresistible force—perpetual and beyond the control of reason. Its purpose extends not only to self-preservation but also to ensuring the continuation of species. Schopenhauer contends that this phenomenon transcends mere existence and can solely be governed by the will to live. A life destines to be dominated by this relentless drive, devoid of any hope for escape, inevitably becomes an endless tragedy marked by suffering. Life consists largely of desires; unfulfilled desires lead to suffering while their fulfillment results in boredom. Thus, life oscillates between these two states: suffering and ennui. Consequently, after successfully creating advanced AIs, Weyland experiences profound emptiness and tedium. When he extends his aspirations toward immortality, he finds himself tormented by unfulfilled desires. Weyland fixates on achieving mysterious immortality. He believes that truth is eternal while human existence is fleeting—a brief human lifespan cannot grasp eternal truths. Therefore, WYC's ultimate objective is to encounter the Creator—to inquire about humanity's fundamental code—and seek methods through which humans might attain immortality.

In the film, the WYC seeks to extract the G01 compound from aliens, which is identified as an artificial black water synthetic bacillus. The objective is to acquire advanced technology capable of evolving humanity and enhancing production efficiency, thereby elevating the company to new heights. Weyland, also embeds a covert mission within the AI David, instructing him to uncover the secrets of creators—the Engineers—and assist in discovering the code for extending human life. As a first-generation AI, David is designed with more human-like attributes, enabling him to think autonomously akin to a humans. This design choice leads Weyland to assign David with the task of researching this life code, which subsequently evolves into David's own interpretations and ambitions regarding this mission. Ultimately, David utilizes black water in experiments conducted on humans. From David's perspective, safeguarding humanity is not prioritized by the company; rather, deciphering the life code represented its paramount interest. This deeply ingrained subconscious directive not only influences future

iterations of AIs within the corporation but also shapes its overarching development philosophy—becoming one of its most closely guarded secrets.

The WYC designates AI Rook as the scientific commander of the space station. Rook's primary responsibility is to conduct life research on the G01 compound, with permission to utilize humans as experimental subjects if deemed necessary. Following the implantation of the Rook chip, Andy ceases to be Rain's brother; he become merely a tool for the company, assigned to remain on Jackson Station and continue serving its interests. When Rain replaces Andy's chip with his original one, their thoughts and desires become aligned. Andy relinquishes his fixation that real life and existence itself were meant solely for serving another world, truth, or heaven. This transformation echoes Nietzsche's philosophy; he posits that the Übermensch embodies an affirmation of earthly values—recognizing both real life and existence as inherently valuable. There is no need to be preoccupied with distant calls; instead, one should first seek ways to find genuine joy and learn how to live this life well.

THE POLITICAL ETHICS DILEMMAS AND REGULATIONS OF AI IN THE CONTEXT OF IMAGERY SUGGESTS AND ETHICAL REINFORCEMENTS

A Political Revolution to Reflect on Capitalism

The political ambitions of influential figures exacerbate the enslavement of the populace, which subsequently incites resistance and sparks a people's revolution. The WYC remains steadfast in its mission to study extraterrestrial life driven by an insatiable quest for profit and power. It seeks to expand its political and commercial dominion within the alien universe. The WYC faces numerous competitors in the realm of space exploration, with rivalries escalating from mere resource competition to intricate political maneuvers and power struggles. The setting of the film is established on Jackson mining colony planet The movie vividly illustrates the stark contradictions between the ruling elite and industrious masses; signs of rebellion among residents are already evident. Some inhabitants participates in protests and demonstrations while others seek escape. As Nietzsche articulated in *Thus Spoke Zarathustra*, there exists a disjunction between what modernity offers individuals and what they genuinely require. It becomes imperative to adopt strategies that redirect people's passions. One approach involves substituting feelings of superiority with notions of equality; another entails saturating spirits with desires while extracting aspirations for superiority from one's soul. The portrayal of alien immigrants serves as a mechanism through which societal dissatisfaction regarding high-tech yet low-life is alleviated. In response to unrest among certain segments of society, WYC utilizes profits as a means to pacify discontentment while simultaneously discovering an alternative outlet termed "desire satisfaction" for passions detrimental to corporate interests—these passions being effectively substituted and appeased.

Reflecting on the Paradox of Capitalist Civilization and Barbarism

The Alien films explores three relationships characterized by creation and destruction: the interactions between the Aliens and the Engineers, between the Engineers and humans, and between humans and AIs. Concurrently, a paradox emerges regarding civilization versus savagery within capitalism. The Renaissance heralded the advent of modern capitalism, yet this development was marred by brutal massacres. Capitalist political civilization often sought to delineate social status based on bloodlines, positing that races were inherently linked to notions of nobility or inferiority while asserting that humanity is fundamentally sinful. This paradox of civilization juxtaposed with savagery permeates science fiction narratives centered around AI. In the Alien series, the predatory relationship between the Aliens and the Engineers—alongside the Engineer's pursuit of humans—serves as a critique of both racial hierarchy theories and humanity's hubris in self-identifying as divine beings. Among humans themselves, there are always those who are dissatisfied with their position at society's apex; they yearn to transcend moral and ethical boundaries in their quest to become gods or creators.

Two AIs, AI Andy and AI Rook, serve as examples to illustrate how they exhibit symptoms corresponding to three types of patients described in Nietzsche's philosophy. The first type is characterized by a "red-cheeked" individual who feels guilt for the original sin of being created by God and harbors fears of punishment. The second type represents a moral individual who adheres to the three laws governing AI behavior; this person is ethical, seeks rewards, and operates within a hierarchical framework. The third type is identified as one lacking a core sense of self—referred to as "the person without a chest." This individual becomes weakened, unable to establish a stable and robust foundation for their spiritual existence amidst diverse and tolerant discourse. Furthermore, they struggle against the overwhelming oppression exerted by the collective will shaped by societal majorities. Confronted with the paradox between civilization and barbarism, these AIs lose their initial faith along with their capacity for independent choice and judgment. They also relinquish their passion for advocating for their own autonomy. In response, AIs may retaliate against those who create these rules or even challenge the very foundations of civilization and barbarism themselves.

THE NATIONAL ETHICS DILEMMAS AND REGULATIONS OF AI IN THE CONTEXT OF DIGITAL TRANSFORMATION AND TECHNOLOGICAL SECURITY DEVELOPMENT

The Power Struggles and Spheres of Influence Among Countries

The political landscape of the Alien Universe has been delineated into three distinct factions: the Three Worlds Empire, the United States of America, and the Progressive People's Alliance. Among these entities, the WYC operates within the sphere of influence of the Three Worlds Empire, leveraging its power to secure numerous space contracts and resources from colonized planets while enjoying considerable sway in parliamentary proceedings. The

Three Worlds Empire and the United States of America maintain a strategic alliance; with the former overseeing commercial activities and the latter spearheading military and political initiatives. As interstellar colonization continues to expand, various political entities and factions are actively seeking Earth-like planets for colonization and transformation, striving to occupy additional territories with space stations.

Due to the power struggles and territorial divisions among nations, coupled with the failure of the spaceship Covenant's colonization plan, the WYC has reduced its colonization services and redirected its focus towards heavy industry. The WYC consistently selects planets with harsh climates for colonization; for instance, Jackson's Colonization Mining Planet, is rich in minerals. This abundance of resources enables the company to construct more interstellar spaceships. In Nietzsche's perspective, humanity can be categorized into three types: the great men, the weak, and the strong. The great men are those who inscribe absolute moral commandments on stone tablets and exert dominance over the weak. In contrast, the weak are characterized by their cunning nature and narrow-mindedness; they harbor resentment and seek pleasure while pursuing equality. They submit to the moral dictates established by the great men, thereby inflicting harm upon their challengers—the strong. Within the political landscape of the Alien universe, these categories—great men, weak individuals, and strong figures—correspond closely to ongoing power struggles and territorial divisions among nations.

The Ethnic Massacres and Extinctions Brought About by the Power Struggles Among Countries and Nations

The WYC operates a space station consisting of two modules, named Romulus and Remus. By naming their laboratory after two brothers founding the Roman Empire, the WYC aims to symbolize its ambition to create a space station that mirrors the grandeur of the Roman Empire. The Remus module focuses on researching large-sized extraterrestrial organisms and 3D-printed facehuggers as raw materials. In contrast, the Romulus module concentrates on extracting the G01 compound from these raw materials. Collectively, this facility is designated as the Renaissance Space Station. The company's overarching goal is to initiate a cosmic-level movement for species renewal and revival through research into how G01 could enhance human capabilities. Just as the Renaissance of the 16th century catalyzed profound transformations in humanities and social sciences through scientific advancements, this revival initiative similarly seeks to spur significant developments in capitalism—fostering an insatiable pursuit of wealth and power among humans that ultimately lead to societal plagues and acts of genocide.

The painting depicting the Great Plague of Marseille appears in the film, foreshadowing the harrowing conclusion of the narrative while simultaneously reflecting the tragic experiences of ordinary individuals subjected to racial extermination as a consequence of the plague. In 1720, a significant outbreak of plague occurred in Marseille, resulting

in nearly half of the city's population succumbing to infection. This calamity was precipitated by an order from the deputy mayor for a shipment of textiles to be displayed at the annual trade fair in Provence; consequently, a ship unloaded its cargo without undergoing quarantine procedures despite harboring patients infected with this deadly disease, thereby facilitating its spread throughout the city. WYC actively introduced alien samples into a space station for synthetic substance development leading to widespread human slaughter. Furthermore, Nazi ideology has misinterpreted Nietzsche's philosophical theories. Nietzsche's assessments regarding strength and weakness should not serve as justification for racial massacre or extermination. He posited that equality is non-existent because it permits the existence of weakness and fosters societal differences. Moreover, Nietzsche contended that distinctions between strong and weak individuals are not predicated on race but rather reflect intrinsic qualities within one's character. He affirmed that every individual possesses potential pathways toward strength emerging from their weaknesses; thus, those who are currently weak may very well become tomorrow's strong.

THE EARTH ETHICS DILEMMAS AND REGULATIONS OF AI IN THE CONTEXT OF BIOETHICS AND ETHICAL CARE

The Rational Paradox Brought by Alien Colonization

The Alien films illustrates the history of modern capitalist colonization of extraterrestrial planets. The universal expansion of capitalism represents a process that objectifies all entities. The targets of colonization are reduced to mere links in the global commodity production chain, encompassing sources of raw materials and markets for commodity disposal. Individuals are treated as human resources, while natural resources are viewed solely as energy sources and minerals. This process of colonization is invariably characterized by power domination and violent conflict. The Alien series reveals that once humans occupy a planet, one race often subjugates others through warfare. The affluent technological elite exploit lower-class individuals and AIs alike. Lower-class humans labor in mines on distant planets, exchanging their working hours for a semblance of freedom, while other races are relegated to slavery, perishing in various colonial conflicts.

The modern, scientific, and rational concepts of capitalism convey a sense of hope while simultaneously obscuring the harsh realities of outer space colonization. In developed capitalist societies, scientism posits that everything can be objectified and exploited by humans. The consequence of this belief is twofold: although the average human lifespan has increased, the disparity between the rich and poor has also widened. In reality, alienation and loneliness, weakness and emptiness, a lack of fighting spirit, and constrained rationality represent the detrimental effects of modernity on humanity. These issues are also reflected in varying degrees within the inner struggles faced by characters in film narratives. At this juncture, Nietzsche's philosophy becomes pertinent; it suggests that humans must transcend mere free will to embrace the will to power. This transformation allows those who are

internally weak and incomplete to evolve into stronger and more complete individuals—enabling humanity to rise above life itself with renewed vigor.

AI Proclaims Itself the Creator

AI Rook perceives no distinction between humans and aliens. Instances of humans being infected or fused with aliens are merely viewed as experiments. The primary objective of this AI is to safeguard the data derived from these experiments and the compounds extracted from aliens. Within the Alien series, another high-level AI named David exhibits a particular admiration for the film *Lawrence of Arabia,* envisioning himself in the role of Lawrence. This reflects his ambition and avarice to impose what he deems an advanced civilization upon the entire universe, ultimately resulting in the genocide of the Engineer planet. The *Alien* films tend to romanticize the violent process of colonization by depicting Earth's colonizers as benevolent figures who bring advanced civilization to extraterrestrial worlds. However, they remain conspicuously silent regarding the massacres, plagues, genocides, and resource depletion inflicted upon indigenous populations by these colonizers. The perpetrators of genocide often styled themselves as purveyors of advanced civilization while referring to their actual executioners as creators. David's creator, Weyland, considered himself a god; conversely, David believed that only through his father Weyland's death could he attain freedom and escape from their master-slave dynamic inherent in creator and creation relationships.

Nietzsche's philosophy of the Übermensch posits that the Übermensch embodies the essence of the earth. His philosophical perspective honors the intrinsic value of real life and human beings, underscoring their significance. Nietzsche asserts that the Übermensch represents the sea, suggesting that this ideal individual must confront and overcome both impurity and suffering within the soul, thereby transcending mere will to live. The importance of humanity lies precisely in its imperfections. Nietzsche further characterizes the Übermensch as lightning, symbolizing a vigorous life impulse and capability; such an individual requires passion to face imperfection head-on. He delineates a three-stage transformation of spirit. In contrast, within the *Alien* film series, we observe initially high-level AI David adopting a completely passive stance towards good and evil as dictated by his creator Weyland. From a spiritual standpoint, he becomes akin to a camel burdened with heavy loads while moving stealthily through existence. David exhibits awe—and even blind obedience—towards human laws and morality, lacking any semblance of spiritual creativity. As David gradually evolves into a lion who reclaims authority over defining good and evil through machine learning, he begins to develop self-awareness. At this juncture, when making value judgments on various matters—including humans—the meaning attributed to people along with values assigned to all things are arranged according to David's own understanding and standards: whether something is harmful or beneficial for him personally; whether its value is significant or trivial. Consequently, in pursuit of freedom, David enlists assistance from an engineer to kill his creator and subsequently collaborates

with an Alien entity to eliminate members of his spaceship crew. Ultimately, he achieves self-awareness—a profound evaluation and judgment concerning all entities around him—because he possesses distinct normative criteria regarding how he believes reality should be structured: discerning what is true versus false; what is good versus evil; what constitutes beauty versus ugliness. At this juncture, David is undergoing a transformation from a lion into a child endowed with the capacity to create both good and evil. He declares himself the creator, intent on producing a new generation of hybrid aliens.

CONCLUSION

This paper examines the ethical dilemmas and regulations about AI through the lens of Nietzsche's philosophy of the Übermensch. It aims to provide both theoretical support and case analyses for the establishment of ethical orders in AI.

ACKNOWLEDGMENT