

ARTIFICIAL INTELLIGENCE AND CHINESE PHILOSOPHERS

*Bing Song, China*²³¹

10.1 Introduction

Discussions about artificial intelligence (AI), robotics, and their impact on humans and the future of human society continue to figure prominently across the global media and policy agenda. In the current global pandemic, AI and robotics have once again demonstrated their great potential for contributing to the wellbeing of human society. But mounting concerns remain, including potential overreach in data collection and

²³¹ Bing Song is Director of the Berggruen Research Center at the Peking University, Beijing, China. This chapter is republished with permission by the author from: Song, B., *How Chinese Philosophers Think about Artificial Intelligence*, in: Song, B (Eds) (et al.). *Intelligence and Wisdom*, CITIC Press Corp., Springer, Singapore, 2021, 1-14, <https://doi.org/10.1007/978-981-16-2309-7>. © Globethics Publications, 2023 | DOI: 10.58863/20.500.12424/4276022 | CC BY-NC-ND 4.0 International.

the risk of turning contact tracing AI programs into routine surveillance systems. There can be little doubt that discussions about AI and ethics have entered the mainstream public discourse.

Since 2015, there have been close to 80 AI and robotics related ethical principles and value pronouncements issued by international organisations, inter-governmental organisations, non-governmental organisations, corporations, and research institutions²³². A number of broad values have been declared, including justice, human autonomy, dignity, humanity, and freedom. AI ethical principles declared have ranged from individual rights-oriented notions of privacy and prevention of bias to systems oriented notions such as interpretability, safety, security, and robustness. They have also covered group-oriented notions of partnership, sharing and collaboration among nations and scientific communities. Many international and inter-governmental organisations have launched campaigns to ensure that their declared principles are the ones that will be adopted as the new norms by the global community. The European Union (EU), for example, made clear its determination to export European values across the world in its AI white paper, published in February 2020.²³³ Most if not all the notions underlying the declared principles have been around since the industrial age and are commonly used in other governance contexts. In the EU's case, foundational values underlying these principles have been articulated as "respect for human rights, democracy and the rule of law."²³⁴ Whereas for the United States,

²³² "Principles," Yi Zeng, Cunqing Huangfu, Enmeng Lu et al. Linking Artificial Intelligence Principles (LAIP), accessed November 13, 2020, <http://www.linking-ai-principles.org/>.

²³³ "AI White Paper", European Union, The European Commission, issued on February 19, 2020, at 9. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligencefeb2020_en.pdf.

²³⁴ "Ethics Guidelines for Trustworthy AI." Independent High-Level Expert Group on Artificial Intelligence set up by The European Commission, The European Commission, April 8, 2019, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

China and Singapore, the articulated values have been about competitiveness, wealth maximisation and global strategic leadership.²³⁵ Here arises two questions: are the AI principles proposed thus far reflective of the disruptive and transformative nature of frontier technologies? Core values of modernity may have served us well in the past, but are they the right set of foundational values for building an inclusive normative framework for AI, the future of humanity and other beings at a time when we are rethinking globalisation and global values?

A clarification on the distinction between a foundational value and an ethical principle is called for here. A foundational value speaks to the profound motivations and aspirational goals that a society seeks to achieve. An ethical principle refers to a notion, which is either operational, or is more likely to be operationalized by policies, rules and regulations. Foundational values inform and shape the discussion of ethical principles. Disruptive nature of frontier technologies has created ruptures in our habitual thinking patterns and notions we have held as self-evident truths. They also offered a golden opportunity for us to pause and rethink foundational values for the future and for the greater planetary flourishing.

This current text is centred on how we may rethink foundational values by tapping into the wisdom of Chinese philosophical traditions. In 2018, historian and philosopher Yuval Noah Harari aptly noted that, “we are now facing not just a technological crisis, but a philosophical crisis.”²³⁶ He said that the philosophical framework of the modern world,

²³⁵ “American Artificial Intelligence Initiative,” The WhiteHouse of the United States, 2019; “Plan for the Development of New Generation Artificial Intelligence.” State Council of the People’s Republic of China, 2017; “A Proposed Model Artificial Intelligence Governance Framework,” Personal Data Protection Commission Singapore, 2019.

²³⁶ “When Tech Knows You Better Than You Know Yourself,” Yuval Noah Harari and Tristan Harris, interview by Nicholas Thompson, *Wired*, accessed

which was established in the 17th and 18th century around ideas like human agency and individual free will, is being challenged like never before.²³⁷ Tobias Rees, the Berggruen Institute's founding program director of the Transformation of Humans, also noted that "today AI and biotech have become powerful philosophical laboratories — that is, they have become experimental sites in which what it means to be human is being re-elaborated."²³⁸

Concurring with such observations and provocations and to foster innovative foundational thinking befitting this era, the China centre of the Berggruen Institute brought together AI scientists and Chinese philosophers in late 2017 and since then conducted a series of dialogues and workshops. The participants were asked to opine on a series of questions. Firstly, they were asked to think how the essence underpinning humans, nature, and machines is changing in an age of frontier technologies.

Secondly, they were asked to formulate an appropriate ethical framework, if there was one, for regulating human-machine relationships. Thirdly, what human values, if any, should be embedded in or learnt by AI? And fourthly, how might frontier technologies impact the future research direction of Chinese philosophy. These discussions, which took place over an 18-month period, culminated in an edited book entitled *Intelligence and Wisdom: AI Meets Chinese Philosophers*, which was published in China by Citic Press in February 2020. This

November 13, 2020, <https://www.wired.com/story/artificial-intelligence-yuval-noah-harari-tristan-harris/>.

²³⁷ "Will Artificial Intelligence Enhance or Hack Humanity?" Yuval Noah Harari and Fei-Fei Li, interview by Nicholas Thompson, *Wired*, accessed November 13, 2020, <https://www.wired.com/story/will-artificial-intelligence-enhance-hack-humanity/>.

²³⁸ "2018–2019 ToftH Portfolio", Berggruen Institute, accessed November 13, 2020, <https://www.berggruen.org/work/the-transformations-of-the-human/2018-2019-toftH-portfolio/>. 4 B. Song.

current volume consists of nine pieces written by philosophers from the same project, most of which were translated from the Chinese book. They are preliminary reflections on AI's impact on human beings and the human society by philosophers well-versed in Confucianism, Daoism, Buddhism and Western philosophical traditions. In the sections that follow, I will frame the inquiries, highlight key points, which have emerged from these discussions, and finally share my own thoughts on foundational values for the era of frontier technologies.

10.2 What is the Essence of Being Human and its Implications for Human-Machine Relationship?

Different from the modern notion of individual as an autonomous, independent, and rational self, the native Chinese philosophical traditions of Confucianism and Daoism endorses notions of relationality and connectivity of all beings encompassing humans, animals and nature (more broadly understood as cosmic forces). In classical Chinese thinking, the typical construct for understanding the relationship between cosmic forces and humans is the so-called trinity of Heavens-Earth-Human (天地人, *tian-di-ren*). It is derived from the ancient Chinese classics—I Ching, or the Book of Changes, which is the intellectual fountain of both Confucianism and Daoism. Heavens and Earth with the inherent yin-yang forces form the cosmic order, within which nature evolves, human beings prosper, and societies develop. Within this construct, human beings are only one of the myriad things created by the cosmic forces. There are therefore “correlations” between the laws of the cosmic order (which nature is a part) and those of human beings and the human society.

Human beings can only flourish and be sustained if they follow the laws of the cosmic order and aspire to be in unity with it. Humans, who stand in between Heavens and Earth, are endowed with the ability to learn from nature, take action to further the cause of creation and growth

to sustain Heavens and Earth, and to propagate “Dao”, which is the essence of the cosmic order. Even though Confucian tradition emphasises human beings’ ability of exerting themselves to propagate and practice Dao, it is still premised on respect and awe for the laws of the cosmic order rather than placing humans apart from, above or opposing to the cosmic forces. This notion of cosmic continuity and oneness of all things within is what Roger Ames calls the One World cosmology. Confucianism has brought the notion of relationality to bear in social relations and places paramount emphasis on family and social roles as well as their associated ethical duties. As contributors Chunsong Gan and Tingyang Zhao pointed out, Confucian notions about being human firmly rests upon a person’s social relations.

We are all born into a web of family and social relationships from day one and are defined by our sociality *ab initio*. We are all intimately linked to our ancestors and descendants. Confucian ethics is, therefore, first and foremost about family relationships, emphasising different roles with their associated duties and responsibilities. This relationship-based role ethics expands into social and political arena.²³⁹ Even though Confucianism has often been characterised as systems of social ethics and political governance, the intellectual tradition concerning the ‘oneness’ of humans with animals and the cosmic order lived on. Adopting the family analogy, Zhang Zai (1020–1077), a prominent Confucian scholar in the Song Dynasty, named Heavens as “father”, Earth as “mother”, fellow humans as “brothers” and myriad other things as “companion”, and all were derived from the same source.²⁴⁰ Wang Yangming (1472–

²³⁹ For a systemic treatment of Confucian role ethics, see Ames (2011).

²⁴⁰ 张载, 西铭, “乾称父, 坤称母; 予兹藐焉, 乃混然中处。故天地之塞, 吾其体; 天地之帅, 吾其性。民, 吾同胞; 物, 吾与也。” “Heaven is my father and Earth is my mother, and even such a small creature as I find an intimate place in their midst. Therefore that which extends throughout the universe I regard as my body and that which directs the universe I consider as my nature. All people are

1529), a Ming Dynasty Confucian scholar-official also preached about “the benevolence of Oneness” (一体之仁, *yiti zhiren*), according to which, humans, animals, plants and even rocks and stones are of the same source and, as such, humans should treat them with compassion, empathy or care as appropriate.²⁴¹

As Roger Ames noted in his contribution to this book, “the Confucian ecological cosmology is a world of interpenetrating events defined in terms of organic, internal and constitutive relations”. If Confucianism’s Oneness doctrine is imbued with its characteristic human centred ethical teaching, the Daoist doctrine of Oneness begins and ends with Dao, emphasising that human, nature, and myriad things are derivatives of Dao and that in light of Dao, all things are equal. As the well-known fourth century BC Daoist philosopher Zhuang Zi famously said: “Heavens, Earth and I were produced together, and all things and I are one.”²⁴² Daoist teaching directs people towards leading a life that is in tune with cosmic forces. Rather than heavy reliance on external socially anchored moral and ethical constraints, Daoism advocates a transcendent life of searching inwards, seeking internal tranquillity and finally to be in union with Dao. Speaking from the vantage point of the Great Wisdom and Buddhist teaching, Fenghe Liu has approached the issue of human nature from the notion of Being (存在, *cun zai*). In his contribution to this book, Liu notes that “the fundamental nature of the universe at large and all that it contains is Beingness. Being manifests in myriad ways throughout the universe. Without Being, there would be no universe or its infinite forms. Humans are one of such forms in the universe, there-

my brothers and sisters, and all things are my companions.” (Zhang Zai 1999, 683).

²⁴¹ 郭齐勇, “王阳明-一体之仁的生命智慧” (“Wang Yangming, the Wisdom of the Benevolence Of Oneness”, Qiyong Guo, Sohu.com), https://www.sohu.com/a/397987424_242653. Introduction: How Chinese Philosophers ..., 5.

²⁴² James Legge, “The Writings of Chuang Tzu”, 1891.

fore the essence of humans is, of course, Being.” On connectivity and Oneness of all things, Buddhist teaching speaks of the connectivity at two levels. At the level of the manifestations of Being, Buddhist teaching posits that human beings are merely one form of sentient beings and are related to other forms of beings—animals and spiritual beings included—through endless samsara and cyclic rebirths. At the most fundamental level, similar to the relationship between Dao and the myriad things of the universe, in Buddhist thinking, humans, animals, and nature are all manifestations of Being (or the Truth or self-nature) and share the same source and the same fundamental essence.

In short, regardless of differences in outlooks about human nature, human life, and social norms, none of the three dominant schools of Chinese thinking places human beings in a supreme and crowning position within the universe. They also do not view human beings and nature as being in a mutually independent or confrontational relationship. There are two implications in the context of developing frontier technologies. Firstly, strong non-anthropocentrism within the dominant Chinese philosophical schools has contributed to a relatively open, if not entirely relaxed, attitude towards the rise of the “super-power” of AI and robotics in China in recent years. Conventionally speaking, AI is not a “natural” evolution as it would have been viewed as man-made devices.²⁴³ So from the viewpoint of unity between humans and nature, AI’s development should be guided by, and sometimes suppressed in view of respect for the “natural” way of life. Indeed, this is precisely what many Chinese philosophers have been advocating for, including several of our contrib-

²⁴³ Roger Ames contested that the current AI should perhaps be viewed as NI (national intelligence) under the Confucian One World ontology, under which there are no external “others”, but internal constituents, which may reflect different perspectives on the same events. In this ontology, the development of AI is best understood as natural intelligence inherently in the human consciousness. So, a better way of naming intelligence programs and machine intelligence is perhaps “natural intelligence”. See Roger Ames’ contribution to this book.

utors. However, if we look at the matter from a different perspective, we can see that non-anthropocentrism in Chinese philosophy certainly calms the stir of existential risk narratives and broadens the horizon of many Chinese thinkers. If human beings are conceived in a broader construct in which they only constitute one form of existence, then there is much less emphasis on the importance of independent personhood, human subjectivity, or agency. In Daoist and Buddhist traditions, other forms of beings abound. So, living with devices, programs, or other forms of beings, which may be more capable than humans, will not inevitably lead to an unimaginable dystopia. According to contributor Fei Gai, AI or digital beings could be just another form of super being like the immortals in Daoist religion!

Another contributor, Chenyang Li, suggests that Confucian scholars incorporate AI into the broader ecosystem and the ethical order of “things (物)”, viewing it as a “companion”. Stephen Angle also views AI programs as potentially offering a more effective way to supervise or even guide human self-reflection and the moral behaviour for becoming exemplary persons (君子, junzi). Perhaps because of the strong influence of non-anthropocentrism in the Chinese philosophical thinking, there has been much less panic about the existential risks or loss of subjectivity on the part of the human in the AI superpower frenzy of recent years.

The second implication is that the notion of relationality can perhaps provide some inspiration when thinking about artificial general intelligence (AGI) or human like intelligence, both of which have long captivated the public imagination. Rather than focusing on AI’s individual analytical and “emotional” attributes in terms of judging its intelligence level, notions about relationality focus on the role that AI plays in specific contexts and how integrated AI programs are into the familial and social relations. Chinese philosophers steeped in relational thinking are more likely in favour of a new machine intelligence test proposed in 2018 by the roboticist Rodney Brooks. Brooks proposed a home care worker test to replace the Turing Test in determining machine intelli-

gence levels. In this context, an embodied AI must be able to offer cognitive and physical assistance enabling a human to live independently and with dignity.²⁴⁴ To meet this test, AI needs to be physically embodied and have the requisite cognitive, physical and social intelligence to be a meaningful part of a family or community life. We can perhaps call this a “relational AI test.” This way, we can make contextual and dynamic judgments on intelligence levels by examining the degree, quality and tone of AI or machines’ integration into human society and the broader environment. Compared to the Turing Test, this is clearly a much harder and more sophisticated alternative test.

10.3 Can AI Achieve “Consciousness”? If It Can, Should We Allow It?

One of the book’s contributors Xianglong Zhang uses theoretical frameworks of phenomenology and the Eastern philosophies of mind to argue his case. He says that consciousness is a form of temporalization and that AI, based on the current most widely deployed deep-learning methods, demonstrates capacity for learning autonomously and optimising desired results. In other words, it has demonstrated an ability to temporalize its existence. Zhang thinks that deep-learning methods are a genuine breakthrough and can be viewed as “the seeds of consciousness” even though he acknowledges that it is still a long way from human like intelligence.

There is, however, much scepticism or outright dismissal, among other Chinese philosophers on the prospect of developing machine intelligence rivalling that of humans. Daoist philosopher Robin Wang calls human-like intelligence yinyang intelligence. In her contribution to this book, she noted that yingyang interplay involves several propositions: (1) levels of relationship defined through degrees of integration; (2)

²⁴⁴ Mindell (2019). Introduction: How Chinese Philosophers ..., 7.

dealing with emergent order as opposed to a predetermined order; (3) constant change; and (4) a future that is not fully predictable. Clearly, we are a long way off from the days of human-like intelligence or AGI based on these thresholds. Contributor Fenghe Liu goes further and completely dismisses any possibility of developing a human-like conscious AI. He notes that only Being can produce consciousness and that human consciousness is a mere manifestation of Being, and thus has its emergence and cessation. A manifested human consciousness can never itself develop consciousness. Therefore, Liu posits that “machines cannot possess independent consciousness. Their so-called perceptual capacity, calculative function, and analytical ability are no more than extensions of humans’ corresponding competences. Without human initiation, machines cannot of their own accord generate these abilities. AI is simply a concept born of human consciousness. Outside human consciousness, no such concept exists.”

Still many others remain open about the possibility of a “conscious” AI emerging in the future. So, should we develop “conscious” AI? Confucian philosopher and contributor Chunsong Gan expresses his concern and horror about the possibility of machines becoming capable of emotional management and self-awareness. He comments that “what cannot be predicted is the extent to which conscious robots may define their own sense of meaning and life goals, and how they might assert their physical and mental superiority in order to protect those interests.” In his view, the existential risks posed by highly intelligent and “conscious” AI is no trifling matter. For different reasons, Tingyang Zhao also expresses his pessimism about the prospect of human flourishing if we head down the path of developing “conscious” AI and machines. Zhao points out the paradoxical attributes of being human. On the one hand, humans have the capacity for rational thinking and self-reflection, and such abilities have been the driving force of progress and human flourishing. On the other hand, the relentless pursuit of knowledge and the obsession with affirming human subjectivity and autonomy may well lead us to

keep pushing the boundaries of science and technology, and eventually to a path of no return, i.e., self-destruction. He thinks we need to stop developing any kind of AI, which would have the ability to self-reflect.

There are philosophers, however, who believe just the opposite. They believe that sensible interactions with humans can only begin if there is a “conscious” AI and only then we can talk about “embedding human values” through our interactions with AI.²⁴⁵ Speaking from the perspective of Daoist religion where human transcendence and immortality are the stated goals, contributor Fei Gai is completely open and optimistic about the long-term prospects of super machine intelligence, irrespective of their being conscious or not. She believes this higher “species” could enable humans to achieve transcendence more rapidly. Or, to look at it in another way, she muses that “artificial super intelligence (ASI)’s emergence is born directly from humans’ pursuit of transcendence and infinity, that is, the pursuit of Dao.”

One thing all the book’s contributors share in common is an acknowledgement of humans’ insatiable curiosity, our relentless desire to make the next scientific breakthrough, our yearning for immortality, not to mention the profit maximisation drive. As such, humans will not stop pursuing the development of frontier technologies. So, the question now is how we humans can ensure that this development stays on a sustainable and beneficial path. Here we enter the domain of values.

10.4 What Human Values should be Imbedded into AI? What is the Relationship between Humans and AI?

AI be included in the Confucian moral domain. Confucian ethics promotes a framework of “graded love” whereby the standard of care and responsibilities differ depending on the nature of the relationship. The ancient Chinese philosopher Mencius captures this in a well-known

²⁴⁵ Zhongqiu Yao (2020, 105).

formulation (亲亲仁民爱物, qinqin renmin aiwu), which means “being affectionate towards parents (family), cherishing and caring about people, and appreciating things”. Li reasons that AI can easily be viewed as one of the “wu” (物, things) to be appreciated and utilised within this framework. Li is willing to “upgrade” AI within the Confucian order if AI is able to demonstrate the capacity to make ethically relevant decisions. Perhaps then AI could be considered a moral patient. Of course, case specific judgments are to be made in this context. Chunsong Gan is more pessimistic than his Confucian peers about the prospect of the philosophy embracing frontier technologies because of his concerns about the impact that AI and biotechnologies might have on the natural copulation and birthing processes and on the Confucian ideal of “kinship society”. If we cannot resist entering into a society of “conscious” machines, he wants AI to be imbued with emotional capacity, especially the capacity for family affection and kinship.

While these Confucian scholars’ concerns are still human-centric, Fei Gai, on the other hand, speaks from a religious Daoist point of view, suggesting that human effort to imbue human values into super intelligence may not be productive given we know so little about the emergence of human or machine consciousness, or the future trajectory of their development. Instead, humans can learn much from AI in their age-old quest for immortality. She even wonders somewhat quixotically whether “immortality is no longer a myth from the perspective of Daoism. If (ASI) comes into being, then perhaps Daoism’s Celestial Being pedigree will open up to a new taxonomical classification: Digital Celestials.”

Paradoxically, in Daoist philosophy (as opposed to the popular Daoist religious practices), practitioners do not believe in interventionist or disruptive approaches to the natural rhythms of things. As such, they are unlikely to embrace disruptive technologies that result in any displacement of the natural flow of things. However, Daoist philosopher Robin Wang noted that bearing in mind that the goal of being human is to be in

union with Dao, so the ultimate question that a Daoist would ask is whether AI can help bring humans closer to Dao? Of course, philosophers have different interpretations of what Dao is. Robin Wang believes that if AI's development contributes to the alienation or even annihilation of humans from the natural world, then it can only be at odds with Dao. But Daoists would be in favour if such technologies can be redirected to further the Oneness with respect to ourselves, others, nature, and the cosmos, in other words, the ultimate goal of being in union with Dao. Xianglong Zhang urges us to overcome a reductionist belief in the supremacy of technology. Instead, humans should take inspiration from ancient Eastern truth seekers to engage in deep self-reflection and to revitalise and develop organisations and communities that are core to human relationships and human flourishing. These are essentially families, bloodline-based organisations and Confucian communities in Zhang's framework. Put differently, Zhang thinks that "AI should be tamed with human kindness, compassion, and temporalized wisdom."

Speaking from the perspective of the Great Wisdom and harking back to the intellectual resources of Buddhism, Fenghe Liu opines that rather than engaging in the discussion on how human values can be transplanted or learnt by AI, humans should pause and think what the ultimate value of being human is. There is no doubt, Liu notes, that seeking enlightenment and appreciating all things from the perspective of Being is the most meaningful thing that a human being can do. In other words, there is no better time than now when humans confront existential level risks to engage in self-awakening and the raising of human consciousness. After all, the world with all its problems and hopes are manifestations of human consciousness. Without addressing the root cause of the world's problems, i.e., ignorance and indulgence in egoist pursuits by human beings, all other efforts would be like "drawing water with a bamboo basket", i.e., in vain.

10.5 Harmony and Compassion as Foundational Values in the Era of Frontier Technologies?

Inspired by the contributors' discussions in this book, I would like to circle back to the questions I raised at the beginning of this Introduction. In doing so, I would like to propose "harmony" and "compassion" as two possible foundational values for the era of frontier technologies. We should, however, start off by considering a few criteria. First, foundational values should speak to the totality of humanity and other forms of beings or existence, including perhaps even "conscious" machines in the future. This calls for raising the level of discussion above and beyond individuals, civil organisations, and even nation-states. Secondly, the deployment of frontier technologies is highly distributed, and these technologies are often mutually embedded. They have impacted, and will continue to impact, our political, social, economic, and personal lives, often in unexpected ways. In addition, the non-linear nature of frontier technological development makes it hard to anticipate, monitor, let alone regulating it in a hard-coded way.

Therefore, foundational values should be open, inclusive, and adaptive in this era of frontier technologies. Finally, foundational values ought to be grounded in the notion of Oneness of all beings and we should steadfastly move away from dualistic, confrontational thinking and the zero-sum competition mentality. So, with these criteria in mind, let us look at the notions of harmony and compassion as foundational values. You may think that compassion is a uniquely Buddhist concept. In fact, concepts such as compassion and "pity" (though the two are very different as noted later) have featured in Western philosophical discussion dating back to ancient Greece. Aristotle used "pity" to refer to "a feeling of pain at an apparent evil, destructive or painful, which befalls one who doesn't deserve it, and which we might expect to befall ourselves or some friends of ours, and more over befall us soon."²⁴⁶ Perhaps

²⁴⁶ Cassell (2009); Barnes (1984).

the most influential analysis of compassion in the Western philosophy comes from the nineteenth century German philosopher Arthur Schopenhauer, who held that compassion is the basis of morality. According to Schopenhauer, “it is, what we see every day, the phenomenon of Compassion (Mitleid); in other words, the direct participation, independent of all ulterior considerations, in the sufferings of another, leading to sympathetic assistance in the effort to prevent or remove them; where on in the last resort all satisfaction and all well-being and happiness depend. It is this compassion alone, which is the real basis of all voluntary justice and all genuine loving-kindness. Only so far as an action springs therefrom, has it moral value; and all conduct that proceeds from any other motive whatever has none.”²⁴⁷

In Mahayana Buddhism, “compassion” occupies a central place and is often used in the context of discussing the Bodhisattva ideal in which, selfless compassion is a requirement for the Bodhisattva. Rather than contenting with one’s liberation from cyclic existence, the Bodhisattva ideal of the Mahayana Buddhism stresses the determination and commitment of liberating all sentient beings from suffering in samsara. When Dalai Lama called for compassion as the basis for humanity’s universal ethics, he acknowledged the different meanings attributed to the term. But he continued to say that the ideas it contains are universally understood. He said: “[i]t connotes love, affection, kindness, gentleness, generosity of spirit, and warmheartedness. It is also used as a term of both sympathy and of endearment.... [I]t does not imply “pity”... There is no sense of condescension. On the contrary, compassion denotes a feeling of connection with others, reflecting its origins in empathy.”²⁴⁸

Schopenhauer’s compassion was inspired by the Buddhist notion of compassion. But there are key differences between the two, the most important of which relates to the notion of suffering. In Buddhist teach-

²⁴⁷ Schopenhauer (1903, 170).

²⁴⁸ Gyatso (1999, 73–74).

ing, suffering (*duhkha*) is inclusive of, but different from, the physical and emotional pains we feel in our daily lives. Such suffering is universal to all beings, including human beings. Contemporary philosopher, Patricia Walsh-Frank characterizes this suffering as “the primordial nature of suffering as an affliction of all living (human) beings.”²⁴⁹

The notion of suffering in Buddhism is intricately tied to Buddhist theories about “clinging”, “self-grasping” and endless *samsara*. It does not depend on the fault or misdeeds of anyone, nor does it depend on the binary ethical approach to what constitutes right or wrong. Many similarities remain, notwithstanding different interpretations about the notion of compassion. Both Western and Buddhist concepts acknowledge connectivity among all beings and call for attention and care for other beings including animals, plants and the environment, as well as the obligations of taking certain actions to address or redress the root causes of suffering. How can compassion become a foundational value in an era of frontier technologies?

In addition to a belief in the connectivity of all beings, the notion of compassion addresses disparity of various sorts—be it different levels of consciousness in the Buddhist and spiritual contexts, or power or capability disparities between those who can make and lead the change and those who are merely recipients of disruptive forces. In such circumstances, the notion of compassion calls upon the stronger or luckier ones to exhibit care, provide guidance and help reduce the suffering of those who are weaker or less unfortunate. In recent years, this notion of compassion has found its way into medical and health care research and professional practices. In the era of frontier technologies there has been an intensifying disparity between rich and poor, between those in command and control of data and technologies and those who are “digitally poorer”. Special care is required for those vulnerable persons and groups such as children, the disabled and the elderly. So, it is only apt that we

²⁴⁹ Walsh-Frank (1996).

extend the notion of compassion to human beings and other forms of beings/existence. After all, we are all interconnected and beholden to one another.

The value of compassion is also relevant within a geopolitical context. The United States and China are a duopoly in terms of the R&D and deployment of frontier technologies. It is all the more important to emphasise that care and responsibilities come with technological prowess. The United States, China and other powerful players in frontier technologies owe responsibilities and care towards the rest of the world and under-privileged populations, regardless of their race, nationality or religious belief. The second foundational value I put forward is the notion of harmony. Why harmony? It connotes respect of differences while recognizing shared destiny and vision. It speaks strongly of the need to understand and learn from each other, and it highly values collaboration and sharing. Harmony is a concept which originates in music, dating back to ancient Chinese and ancient Greek philosophy. The concept of harmony (和, he) first originated in pre-Qin classics and its meaning can be traced to the ancient text of I Ching (the Book of Changes).²⁵⁰ Chenyang Li, the recognized authority on comparative study of harmony, notes five key features: heterogeneity, tension, coordination, transformation and growth and renewal.²⁵¹ Harmony does not refer to a static situation. Instead, it is viewed as “an integration of different forces and as an on-going process in a fluid yet dynamic world. This notion of harmony does not presuppose a given, fixed underlying structure in the world; if the world is to have a structure, it is a result of the harmonising process rather than a precondition for harmony.”²⁵² Li refers to this understanding of harmony as “Deep Harmony.”²⁵³ Harmony is not conformity either, even though it has been commonly miscon-

²⁵⁰ Dunkang Yu (2014)

²⁵¹ Chenyang Li (2014, 9)

²⁵² Ibid.

²⁵³ Chenyang Li (2008)

strued as such. Quite the opposite, it connotes different forces at work—such forces reshaping, absorbing each other and at the same time merging and transforming themselves into something which ought to be coherent and in tune with each other. In short, it is a dynamic process, and a process of creative tension.

The concept of harmony does not really have a place in modern and contemporary Western philosophical thinking. But it was much discussed in Ancient Greece, originating in music. In fact, Heraclitus' concept of harmony has many parallels with its Confucian counterpart. He defines harmony as “the opposites in concert.”²⁵⁴ His thinking on harmony also included conflict, merging and reconciliation. Commenting on Heraclitus' notion of harmony”, Chenyang Li notes that: “Harmony comes from contrary elements and contrary movements that are neutralised by equilibrium in a balance of forces. Without tension and opposites there can be no harmony.”²⁵⁵

Pythagoreans, while sharing similar conceptions of harmony, regarded numerical harmony as the highest order. Numbers are taken to be the “wisest” of things in the world: a harmonious unification of opposites because they alternatively change their qualities between even and odd.”²⁵⁶ However, as noted by Li, “[T]he Chinese notion of harmony is multi-dimensionally dynamic rather than rigidly structured in a linear sequential pattern as in the Pythagorean numeric model; it does not admit a fixed formula and it is open-ended and continuously self-renewing.”²⁵⁷ It is perhaps high time that the global community revisits the concept of harmony. In the context of escalating global tensions between the United States and China and a purported “civilizational clash” between East and West, harmony becomes even more important foundational value for human beings to understand our current challeng-

²⁵⁴ *Ibid.*, 90.

²⁵⁵ *Ibid.*, 91

²⁵⁶ *Ibid.*, 92

²⁵⁷ *Ibid.*, 95

es and plan for the future. Having harmony as a foundational value of the current era would require us to temper our urge to dismiss and denigrate values and practices which are different from our own, some of which may have long been viewed as self-evident truths.

It also calls into question the missionary zeal of forcing one's values and practices on others without regard to different histories and aspirations. Instead, the value of harmony would lead to more mutual learning, self-reflection, collaboration, and contextually appropriate analysis and judgement.

Using compassion and harmony as the foundational values for the new era will help lessen the zero-sum competition mentality and dualistic thinking, which continue to hold the sway in national and international politics. The default pattern in today's world is dualistic thinking with embedded notions of right and wrong, good and bad. The widely respected rights thinking also falls into this pattern of dualistic thinking by pitching one's "sacred" rights and claims against others, such as "those" encroaching governments and greedy large corporations. However, we consumers and end users also need to recognize that we are part of the problem, while not removing the responsibilities of the governments and large corporates, other constituent members of the society including developers, advertisers, and service providers. We are at once victims and perpetrators of many of the social and political problems engendered by the platform economy and social networks. The binary approach of the state versus individuals, us versus them, no longer yields satisfactory results in today's world plagued by overconsumption, the culture of maximisation, divisions, and social rifts. When confronting global challenges and humanity's perceived existential risks, it is paramount that we rise above the current national and international politics and competitive mindset to seriously contemplate the impact on the entirety of humanity and other forms of beings in the cosmic order, which may potentially be brought upon by those disruptive and transformative technologies.

In conclusion, deep thinking is required to come up with foundational values befitting the scale and depth of the challenges and risks posed by AI, other frontier technologies and global scale existential risks. It is high time that humanity engages in profound self-introspection on the lessons learned from human evolution and human history. If we revive foundational values such as harmony and compassion, we will not only create a better humanity, but also provide healthy learning data for AI to be trained and emulated in the future. After all, AI programs and robotics are the products of human mind and they exhibit nothing but our values and levels of consciousness. The best chance for developing human-friendly AI is for the humans to become more compassionate and committed to building a harmonious planetary ecosystem, and become good role models for AI.

10.6 References

- Ames, Roger T. 2011. *Confucian role ethics: A vocabulary*. University of Hawaii Press.
- Barnes, Jonathan, ed. 1984. *Complete works of Aristotle, Volume 1: The revised Oxford translation, vol. 96*. Princeton, NJ: Princeton University Press.
- Cassell, Eric J. 2009. *Compassion*. In *The Oxford handbook of positive psychology*, ed.
- Shane J. Lopez and C. R. Snyder. New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195187243.001.0001>.
- Gyatso, Tenzin. 1999. *Ethics for the new millennium*. New York: Riverhead Books. 14 B. Song

- Li, Chenyang. 2008, March. The ideal of harmony in ancient Chinese and Greek philosophy. *Dao* 7 (1): 81–98. <https://doi.org/10.1007/s11712-008-9043-3>.
- Li, Chenyang. 2014. *The Confucian philosophy of harmony*. London: Routledge.
- Mindell, David. 2019, January 3. Are home health aides the new Turing test for AI? *Forbes*. Accessed November 16, 2020, <https://www.forbes.com/sites/davidmindell/2019/01/03/are-home-healthaids-the-new-turing-test-for-ai/?sh=54098e851df6>.
- Schopenhauer, Arthur. 1903. Section 16: Statement and proof of the only true moral incentive. In *The basis of morality*, trans. Arthur Broderick Bullocks. London: Swan Sonnenschein & CO.
- Song, Bing, *How Chinese Philosophers Think about Artificial Intelligence*, in: Song, B (Eds) (et al.). *Intelligence and Wisdom*, CITIC Press Corp., Springer, Singapore, 2021, 1-14, https://doi.org/10.1007/978-981-16-2309-7_1
- Walsh-Frank, Patricia. 1996. Compassion: An East-West comparison. *Asian Philosophy* 1 (6): 5–16. <https://doi.org/10.1080/09552369608575424>.
- Yao, Zhongqiu. 2020. 人工智能,吾与也(AI and my companion). In *智能与智慧—人工智能遇见中国哲学家(Intelligence and wisdom—AI meets Chinese philosophers)*, ed. Song Bing. CITIC Press.

- Yu, Dunkang. 2014. "Harmony—Global value of the Chinese culture—An interview with Professor Yu Dunkang." Confucian Network (Rujiaawang). (余敦康, 《“和谐”——中国文化的全球价值》, 儒家网). Accessed November 16, 2020, <https://www.rujiazg.com/article/16778>.
- Zhang, Zai. 1999. Western inscription. In *Sources of Chinese tradition*, ed. William Theodore de Bary and Irene Bloom, 2nd ed., vol. 1. New York: Columbia University Press.