

Merciless justice: the dialectic of the universal and the particular in Kantian ethics, competitive games, and *Bhagavad Gītā*

Michael Yudanin

University of Georgia, Athens

Morality's two components

Morality traditionally includes two components: justice and mercy. The first of them, justice, usually refers to fairness, equal treatment, and common rules according to which different cases are to be considered. The second component, that of mercy, is related to benevolence, kindness, forgiveness, and concern for others. It is indeed hard – even though possible, as we will see – to imagine a viable moral system that will embrace one of these components yet reject the other. Empirical research on the nature of morality supports this view.¹ World religions also embrace both aspects of morality: Buddhism, Christianity, Hinduism, Islam, Jainism, and Judaism all talk about justice and mercy – as attributes of God, moral commandments, etc.

Ethical theories that claim one of these two aspects of morality to be primary attempt to derive the other one from it. Kantian ethics can serve here as a classical example. According to Kant, the sole source of morality is pure reason, reason that is cleansed from any empirical content. This idea is first clearly expressed in the *Critique of pure reason* (KrV, AA 06:385 / A800/B828ⁱⁱ), developed further in the *Groundwork for the metaphysics of morals* (see, for example, GMS, AA 04:408) and the *Critique of practical reason* (see, for example, KpV, AA 05:19-20). The morality of mercy is discussed in the *Groundwork*, specifically in relation to beneficence (GMS, AA 04:424-425). Kant's last work on ethics, *Metaphysics of morals*, undertakes the task of providing beneficence with a sound foundation – as I will try to argue, unsuccessfully. Newer examples of moral theories that emphasize one of the aspects of morality as primary are Kohlberg's theory of moral development that focuses on justice (Rest, 1999) and the approach of his student, Carol Gilligan, that stresses the component of mercy (Gilligan, 1993).

The component of justice can be seen as the *universal* component of morality. In fact, justice *is* universality translated into the language of ethics. We can consider justice in a variety of contexts, for example, as distributive or retributive. However, the common denominator of all of the uses of the term 'justice' is that of common rules independent of any particular circumstances. This independence from the particular is essential to justice.

Mercy is usually defined as the opposite of cruelty and malevolence. Yet if we are to analyze its formal features, we will find out that it is essentially preferential and thus particular. Any expression of kindness, forgiveness, helping the needy requires addressing a particular case. The most salient examples of mercy cannot be explained from the standpoint of justice: pardoning someone who deserves punishment, compassion toward the enemy defeated in a just battle, forgiving intentionally caused harm, etc. All the examples above, when made a universal rule, stop being expressions of mercy. If *all* criminals are always pardoned, the act stops being that of forgiveness and becomes a law; it is now a rule rather than an exception that is made out of compassion toward another human being.

Mercy as a formal opposite of justice

The last observation leads us to Kant's deontological paradigm that is based on the principle of universal justice. Kant's categorical imperative requires each and every maxim, or a principle behind certain act, to be considered in terms of its universalization:

act only in accordance with that maxim through which you can at the same time will that it become a universal law (GMS, AA 04:421)ⁱⁱⁱ

In other words, one can act only when the maxims of her action can be considered as necessary as the laws of formal logic or arithmetic. If the maxim-turned-universal-law does not create contradiction, does not make the context of its application meaningless, it passes the test and is allowed to guide behavior that thus becomes morally permissible. If, on the other hand, it fails the test and cannot be universalized without making itself meaningless, the behavior that follows it is immoral.

Fundamentally, Kant equates morality with universality. Only the principle that can be universal without a contradiction is moral; what cannot be considered as universal is not morally permissible. While this might sound counterintuitive at the first glance, this is precisely what gives the categorical imperative the power of moral *law*. Since the contradictions created by impermissible maxims are first and foremost logical flaws, they go contrary to the very principle of our existence as rational beings. This can be demonstrated by using the terms suggested by Wittgenstein in his *Tractatus logico-philosophicus* and *Philosophical investigations*. Reason, with its formal and universal rules, e.g., the principle of non-contradiction, is the "logical space" of our existence (Wittgenstein, 1922/2003, §2.11). Within this logical realm, we have a variety of activities which Wittgenstein deems *language games* (Wittgenstein, 2009, §7ff), each of which has its own rules yet all are under the rules of reason. When the rules of some specific context of human activity are broken, it loses meaning and cannot be rationally pursued further. Kant's second example in the *Groundwork*, that of deceitful promise (GMS, AA 04:422), demonstrates

this principle. The game of promise is meaningful only when the promises are kept. If the maxim of deception for personal gain is universalized, i.e., everybody and everywhere will necessarily lie for personal gain, promising something will become completely meaningless, and so will deception. Same applies to theft: if the maxim “I take another person’s property whenever I want” is universalized, we get a world where everybody takes whatever he wants. As a result, the language game of property, and hence theft as a concept, fall apart: in the world where there is no property that belongs to somebody, there is no theft either.

The examples Kant brings up in the *Groundwork* demonstrate the power of the categorical imperative – the power it has due to its universal nature. Yet the same universal nature gives rise to its limits. As an example, let’s consider forgiveness – a well-respected virtue both in the Eastern and Western traditions. The maxim behind forgiveness cannot be universalized without undermining itself. To see that, let’s think of a clemency granted to a criminal: if we universalize it, it turns into a law, and then we are dealing not with forgiving, i.e., making an exception, but with following a law. We would get similar results if we try to universalize any other expression of mercy.

Why mercy cannot be universalized without ceasing to be mercy? It seems that the reason lies in its essentially particular nature. Mercy is always particular, it is always an exception and cannot be a rule. Helping another always means helping a particular other, or a group of others. Forgiving somebody is *by definition* making an exception. Same applies to liking somebody: it is a specific, essentially preferential attitude, not a universal rule.

In this context the fourth example from Kant’s *Groundwork* is particularly interesting (GMS, 04:423).^{iv} The example talks about helping another human being in need. If we apply the categorical imperative to the situation, we will see that the maxim “Every time I encounter a human being in need, I will help her” is certainly universalizable and hence moral: there is no contradiction in making it a universal law of nature.^v However, the opposite maxim, that of non-helping, e.g., “Every time I see a fellow human in need, I will ignore his suffering,” is as universalizable: we certainly can imagine a society where nobody helps anybody else. While this society might not be a place to live happily, there is no contradiction in thinking it. Kant realizes that and claims that while we have here no logical contradiction, or contradiction in thinking, there is another type of contradiction – a contradiction in willing. In other words, we cannot *will* such a society, as

a will that decided this would conflict with itself, since many cases could occur in which one would need the love and sympathy of others and in which, by such a law of nature

arisen from his own will, he would rob himself of all hope of the assistance he wishes for himself (GMS, 4:423)

This argument, however, does not have the categoricity of the imperative. This is because it relies on the empirically established limitations of human nature rather than on the formal features of reason. This problem does not disappear even if we deem certain needs and limitations essential to human embodiment, as Barbara Herman does (Herman, 1993, and specifically Chapter 3). This is because, differently from the universally valid claims of reason, empirical findings are contingent and require interpretation in order to be comprehended. One can claim, for example, self-reliance to be a goal so noble that any help given by another human being on the ground of beneficence can be perceived as corrupting and hence as something that ought to be rejected. Nietzsche (Nietzsche, 1917) and Ayn Rand (Rand, 1992) can be brought up as examples of arguing for this sort of attitude – and their argument most certainly stands the scrutiny of the universalization test.

Games: justice without mercy

Theoretical problems with reconciling justice and mercy cast serious doubts on the possibility of deriving mercy from justice, i.e., establishing the morality of mercy on a sound formal foundation. However, an important question remains: can there possibly be behavioral maxims that accord with the universal principles of justice yet lack any consideration of mercy? If those are impractical to follow, then the problem remains purely theoretical, just like many would claim that Nietzsche's and Rand's constructs can be intellectually interesting yet not feasible for any actual human society. If, on the other hand, we can bring forth a whole category of behavioral maxims that embody justice yet have no component of mercy, the problem becomes more pressing.

There is such a category of maxims, and even a widespread one. To demonstrate that, let's consider competitive games. Chess is much older than the categorical imperative, and World of Warcraft and other computer games – significantly younger. Both games have a common goal – achieving victory through the destruction of opponent's resources. Any competitive game requires a set of rules that are observed by all players. These rules are necessarily based on the laws of logic – otherwise they would not be comprehensible to rational human beings; same logic that governs universalization. Hence, the categorical imperative must be applicable to any competitive game. And indeed it is: any move in the game should be permitted by its rules, and thus universalizable by definition – all players can make moves according to the rules of the game. Cheating, on the other hand, is prohibited: any violation of the rules, and certainly a known violation done by stealth, is disallowed, since if all players will break rules, there would be no game.

Are the considerations of mercy applicable to competitive games? Should one have mercy for and exercise forgiveness toward an opponent *qua* opponent? The question seems absurd: the end of the game is to defeat the opponent by destroying her resources. Computer games are most instructive in this regard: by using technology, they bring the struggle as close to its human form as the display allows. The avatars bleed, collapse, and otherwise create vivid impressions of pain and suffering using all means of image and sound. And yet mercy still remains outside of the game world.

Why is it so? This question seems to have an obvious answer. Avatars are not humans, they do not feel pain and do not experience suffering, they are never hungry or thirsty, they do not love and do not hate. Mercy is inapplicable to them. It is important to note here that *material objects*, e.g., chess pieces or paintings, do command some moral treatment – the notion of vandalism as morally reprehensible points to it. Yet there can be no mercy toward an entity that is merely *subject to rules*, a virtual, immaterial thing.

This way, competitive games provide us with a behavioral situation where universal justice is present yet mercy is absent.

Two sources of morality

What can we learn from the example of games as a situation where the criterion of justice is applicable, while the criterion of mercy is not? Firstly, it provides us with a class of instances that exemplify the theoretical claim that justice can exist without mercy. Secondly, analyzing the moral aspect of competitive games might lead us to understanding the specific features that allow for the separation of justice and mercy. These features might figure in other, non-gaming contexts where universal principles of justice coexist with the lack of mercy.

It seems that not only Kantian ethics but any ethical system that is based on universal principles is incapable of giving rise to the morality of mercy. As it was shown above, mercy is essentially particular, non-universal, and has specific object. To show mercy is to make an exception. Such exception can be made for everybody whom a certain moral agent meets in the course of her life, if such is her *general* adopted attitude. Yet this does not make the rule behind this attitude *universal*.^{vi} However, the maxim of no mercy can be made a universal law without contradiction, as discussed above in the context of Kant's fourth example in the *Groundwork*. The universal ethics is the formal morality of boundaries and rules, negative morality, as Kant addresses it (see, for example, GMS, AA 04:454 and MS, AA 06:390). On the other hand, the ethics of mercy is the positive morality of an exception from rules.

The sources of the universal morality of justice lay in reason. Yet what are the sources of mercy? The example of competitive games might help us to find them. As it was argued, in the context of the game we confront humans as bearers of reason but not as flesh-and-blood creatures, creatures with physique and psyche. Embodied humans experience pain and pleasure, feel happiness and distress, love and hate; all these aspects of being are inaccessible to reason alone. They are ever individual and cannot be universalized. Thus, the morality of mercy appears to be related to human embodiment, moreover – it seems to be predicated upon such embodiment.

Human essence, then, seems to be determined by two factors: reason and body. Reason enables us to operate with universals, e.g., with logic. Socrates' *Meno* (Plato, 1997, pp870-897) demonstrates this aspect of humanity with supreme clarity. We can refer to this factor as "rational humanity." The second aspect is the "embodied humanity," the one that follows from the embodied nature of humans. This embodiment, essentially contingent and limited, is a necessary condition not only of our existence in the world but also of certain socio-philosophical concepts. Freedom is one of these. Woody (Woody, 1998) successfully argues that the concept of freedom is impossible to have without admitting physical limitations of humans. Freedom is choice, and to make a choice, one has to have access to a number of alternatives. Yet an unlimited, absolute being cannot choose from a number of behavioral options, and thus exercise freedom: *all* possible options are accessible to it, and it can choose *any* one of them. Moreover, such being would not be constrained by a choice once made, and thus such choice loses its meaning *qua* choice in the world, spatio-temporal world. Only a limited being that has access to a limited number of options can choose. Another example, very relevant to the subject of morality, is the concept of ethical duties that spring from the life form, nature, and social station of the moral agent. It is known in Indian philosophy as *svadharma*, duties determined by one's nature, or *svabhāva* (see Gītā XVII:40 and elsewhere, and a thorough review in Chatterjea, 2002, pp111-123). Having a particular level of knowledge or character; being at a particular place in a particular time; belonging to a certain social stratum or profession; having a certain personal history – all these impact on our moral duties. The duty to help people injured in a traffic accident is placed upon those who discovered them rather than upon those who did not; upon the doctors in the hospital to whom the injured were brought for treatment rather than hospital's accountants or other doctors; upon those who are capable to offer help rather than those who know that they might cause more harm than good; etc. All these are particular aspects of our embodied existence as humans.

It makes sense to use similar logic in analyzing mercy. Mercy cannot be applied to an object that is not limited by being physically embodied. Such being would not feel hunger, fear, pain, or distress. Usually, monotheistic deity is brought up as an example of such a being.^{vii} Yet a chess piece or a computer avatar

can also figure as examples – and much more accessible examples at that. A pawn is a pawn not due to its physical characteristics yet by the power of the universal rules of the game that abstract from any particular instantiation of a pawn which can be a piece of wood or plastic, an image on a computer screen, a mathematical representation, or a mental entity in master chess player’s mind. Such are all objects of rules when they are looked upon as objects of rules and not as physical objects, i.e., looked upon from the standpoint of universality. Such are, for example, combatants and non-combatants in the laws of war – abstract legal constructs rather than specific human individuals. On the other hand, limited beings, when considered as such, can be objects of mercy specifically and of particular treatment in general.

Mercy and embodiment

Human embodiment certainly might be the object of mercy, and it is necessary for the object of mercy to be embodied. This, however, does not make it necessary for moral agents to express mercy toward it. Hence, human embodiment of moral objects alone cannot explain the component of mercy in morality, and leaves the question of its sources unresolved.

In order to provide a full answer regarding the sources of the morality of mercy we need to consider not only the object of mercy but also its subject. As it has been demonstrated, reason cannot be a foundation of the morality of mercy. Hence, we are led to look for the origins of subject’s mercifulness in the specifics of human physiology. Mirror neurons and other neuropsychological mechanisms (see, for example, Iacoboni, 2005; Sripada, 2008; Schulte-Rüther, Markowitsch, Fink, & Piefke, 2007) might be considered in this context. Mirror neurons make human subjects register another’s pain through neurophysiological mechanisms upon seeing certain clues, whether they want it or not. Neuropsychological studies are the most recent development in the long tradition of philosophical and psychological research on empathy, which since the nineteenth century has been claiming that we feel another’s pain when we look at a suffering human being, see or even read a description of suffering.

It should be noted that the argumentation provided in the preceding paragraph is not a traditional philosophical one, and that physiological sources of empathy do not provide any support for its philosophical necessity. This is expected: after all, the phenomenon of mercy is essentially contingent and does not follow from reason, as it has been shown.

An attempt to locate the origins of mercy in human neurophysiology leads to another question: if the source of mercy is contingent, can we require it as a part of a system of morality that will compel all human beings? A thorough answer to this question will warrant a separate paper. However, I believe it is possible here to give an outline for an answer. We certainly cannot postulate any specific *contents* of

mercy, yet we can expect it as a category of behavioral response from certain kinds of beings toward certain kinds of beings. In terms of a moral subject, we can require mercy from creatures that are physiologically equipped for it. It would not make sense to require mercy from any rational creature, e.g., from a machine endowed with artificial intelligence: it can play chess and we certainly expect it not to cheat, yet it cannot feel another's pain. Moreover, we would have little basis for requiring mercy from a tiger or an albatross. Yet we can expect it from humans. In terms of the objects of morality, we might want to introduce a hierarchy of creatures based on their ability to feel pain and need, where mammals might figure high and plants – low. This, in an interplay with the physiology-enabled capacity of the subject to show mercy, will determine its degree and direction.

Rules and the justification of evil

If human physiology predisposes us for mercy, why the creature that is equipped with mirror neurons and empathy is capable of torture, guillotine, Gulag and Auschwitz? Capable of claiming that the less-industrious can and must be allowed to starve, and that the normative society should not spend its taxes on rehabilitating drug addicts and convicted criminals? And at the same time – capable of producing Ashoka the Great, Janusz Korczak and Mother Teresa? The issue with such wide spectrum of moral behaviors was most famously expressed by Dostoyevsky's Mitya Karamazov: "Yes, man is broad, too broad, indeed. I'd have him narrower" (Dostoevsky, 1880/2007, p114). This problem exists not only at the individual level but at the level of human civilization first and foremost. As such, it must be addressed by any ethical theory that does not want to lose sight of the actual moral practice and attempts to understand it in order to formulate prescriptions that would be not only valid but also applicable.

Usually the phenomenon of moral inconsistency is explained by invoking the notion of freedom. If we are to understand freedom as the lack of strict determinism in choosing this or that accessible behavioral alternative (see, for example, Woody, 1998), then an individual can opt for good or evil, cruelty or kindness, indifference or compassion in many, if not most, life situations. Yet the possibility of choice alone does not provide an explanation of why this and not that option was pursued. Freedom of choice answers the question "how," yet not the question "why."

Socio-psychological explanations are different. They attempt to explain the choices we make by behavioral determinism. The suggested paradigm views morality as a social, psychological, or cultural phenomenon, and analyzes moral behavior as resulting from a personal history of the individual – e.g., as acquired in the process of socialization; from the combination of societal pressures; or from other similar factors. It is hard to claim that such analysis cannot predict with a high degree of probability the actual behavior of individuals. Yet despite that, it still does not explain the "why" of moral practice.

Deterministic explanations are of a kind that would be produced by an alien who lands in a park frequented by the dog owners, quietly observes the scene for some time, and then concludes that small furry quadruped creatures control the larger bipedal ones, as the latter are led by the former and follow each and every one of their movements to the best of their ability and agility. Can this explanation predict the behavior of dogs and their owners for the next ten or so minutes? Yes, and with almost certain probability. And yet it is obviously wrong. When we are analyzing human moral behavior as similar to a mechanically determined object, we are contradicting the main premise of morality – personal moral choice. Instead, we are assuming strict determinism based neither on intuition nor on sound research, i.e., a research that would not assume determinism first place and thus beg the question. Moreover, the deterministic hypothesis suffers from internal contradictions that do not seem to be easily reconcilable (see, for example, the discussion in Woody, 1998).

Approached philosophically, the problem can be defined as follows. Humans are predisposed to mercy, most probably – physiologically. Yet from time to time actual moral choices made by individuals and groups go contrary to the considerations of mercy. Three factors seem to define the context of this problem: the possibility of individual moral choice, the universal morality of justice, and the particular morality of mercy. All these factors can be found in any known society, whether contemporary or historical – if not *for themselves (für sich)*, then at least *in themselves for us (an sich für uns)* (Hegel, 1807/1977). The approach presented in this paper can provide an explanation for the phenomenon of evil that would accord with these three factors.

Kant's Copernican turn (KrV, AA 03:012 / B xvi) has demonstrated that our point of view, our intuitions have decisive importance for the way we respond to the world around us. The world of an individual is the world of her understanding, information formed by intuitions and ordered by concepts. The former are blind without the latter, and the latter are empty without the former (Ibid., 04:075, / A51/B75). This analysis, further developed by Hegel into the distinction between facts *in themselves (an sich)* and *for themselves/for us (für sich/für uns)*, goes far beyond the trivial claim that we cannot take into consideration the information that is inaccessible to us. A datum, whether it reflects something external to our body or registers an internal feeling, when taken alone, lacks meaning – the intuition through which we acquire it is blind without concepts. Only that datum that has been interpreted by our conceptual apparatus can become an input for decision making.^{viii}

Hence, in order for the considerations of mercy to come into play and to drive our decisions, the inputs of the decision making process should be understood as relevant to such considerations. In other words, the situation should be construed as a situation where embodied humanity is relevant – embodied, and not

only rational. The humanity of the moral object here should be *understood* by the moral subject *qua* humanity. If, on the other hand, the object is understood as lacking essential humanity, the considerations of mercy will be regarded as irrelevant, and only the considerations of justice will remain – as universal, they are always present. Competitive and especially computer games exemplify this situation. As it was argued above, in a game the humanity of the opponent is seen as completely unrelated to the context of the game. When a player is shooting her opponent’s avatar or taking his bishop, she is differentiating between the context of the game and the context of the world inhabited by people.

The word “inhumane” is usually used to denote merciless treatment of fellow humans. Following what has been said above, this word use can be assigned a deeper meaning. In order to elicit inhumane treatment of people by a moral agent, one needs to *de-humanize* his objects, to strip them of their humanity in the eyes of the moral subject who is making ethical decisions, to get him to see fellow humans as similar to avatars in a computer game or as chess pieces – entities to which no considerations of mercy apply.

How can we strip a human object of her humanity in the eyes of a moral subject? If we are to exclude the rare cases of neurophysiological impairments that prevent empathy mechanisms from working normally first place, the most obvious focus will be the conceptual apparatus that interprets the neurophysiological inputs. While certain concepts can be claimed to form spontaneously, without much reflective deliberation, others are a result of deliberate persuasion.^{ix} The latter are formed by appealing to the universal reasoning, to the context of rules – the context that forms the morality of justice. It is possible to *convince* a moral agent that another being, or a class of beings, are devoid of “real” humanity despite its external manifestations that elicit the reaction of mercy. Once convinced, the moral agent will consciously suppress his spontaneous reaction of empathy and even see it as “animism.” Similarly, some of Descartes’ followers *concluded* through reasoning that animals are no more than automata, leading to most odious manifestations of vivisection.^x And it seems that same reasoning can be applied to *homo sapience*.

History has multiple examples of inhumanity. Those that have been carried out on a mass scale, specifically in the twentieth century, were always accompanied by the de-humanization of the victim in the eyes of the perpetrator. Not only the idea of universality as a criterion for the moral quality of behavior was not rejected – it was embraced and suggested as the justification of the atrocities. The most famous example of such conceptually justified de-humanization is the Nazi racial theory. The Jews were to be exterminated because they were an anti-race (*Gegenrasse*; see, for example, Schmitz-Berning, 1998), not humans in the sense in which the moral agents on whom the persuasion had been focused were

conceived to be humans. The theory here is certainly universal – it is applicable to each and every case and does not allow for exceptions at least at the theoretical level. Yet it might well be that the Nazi racial theory is just one instance of the problem that is potentially much wider. It seems that de-humanization is a logical consequence of any reduction of morality to universal rules.

As it has been argued, mercy is essentially particular. Thus, if all factors that determine moral behavior are to be universal, the considerations of mercy will be excluded. For example, deterministic view of history as a class struggle, racial competition, or a survival contest of biological categories does not endorse mercy toward specific individuals. Not only the universal law taken by itself would not prescribe it, it will actually see it as immoral. If universal rules determine everything, there is no room for mercy. Human being turns into an element in the overall scheme of universal laws, similar to a computer avatar or a chess piece, the essence of which is not in its humanity but in its position in the context of rules. Lying and stealing in this context is prohibited – these would break the rules, thus making the context meaningless. Yet helping a stranger in distress without any other reason but kindness would be, in the best case, excluded from the sphere of morality. Mercy would be laughed at, the way Cartesian physiologists mocked the “Pythagorean” confusions of people who pitied animals. In the worst case mercy will be punished, as expressions of compassion toward the victims of political persecution constituted a punishable offense in the USSR and Nazi Germany.

It should be noted that the philosophical argument above accords well with the empirical findings regarding personal ethics and obeying authority. In his groundbreaking experiments, Stanley Milgram found that individuals were willing to cause severe harm to fellow humans while obeying authority, even though the particular authority had little sway over them (Milgram, 1963; see also review at Chapter 6 in Cialdini, 2009). More recently, some research has been conducted that might even point to neuropsychological underpinnings of acquired opinions and their interplay with the core mechanisms of the brain (see, for example, Molnar-Szakacs, Wu, Robles, & Iacoboni, 2007, and review and evaluation at Olson, 2008). What philosophy can contribute here is the “why” explanation. Why do people agree to continue administering electric shocks, close the door of the gas chamber, or switch the TV channel indifferently after hardly glancing over the images of death and destruction? Because mercy, kindness, compassion is taken out of consideration within the scheme of universal rules by turning the suffering human being into a mere element of the universal scheme of things – into an object of experiment, inevitable historical progress, biological necessity, laws of economy, statistics, or religious absolute.

Possible solution: *Bhagavad Gītā* and the complementarity of justice and mercy

Can justice and mercy coexist if they formally oppose each other? While a thorough analysis of this issue goes beyond the scope of the current paper, I would like to provide an outline of a possible answer. It seems that the relationship between justice and mercy is one of dialectical opposition. Hence, they necessitate each other by virtue of opposing one another. Mercy is a formal opposite of justice and is defined through this opposition: it is particular, and as particular it can be comprehended only by being considered as opposing the universal. Justice, on the other hand, is universal and thus defines itself through denying any particularity.

Similarly to Kant's intuitions and concepts in the first *Critique*, mercy without intuition is blind, yet justice without mercy is empty. Kant's universal law is formal and thus devoid of content. It can be "filled" with virtually anything having a character that yield itself to a rule, even with something totally opposing the spirit of Kantian morality, for example, the infamous "categorical imperative of the Third Reich" (Arendt, 2006). On the other hand, mercy without the form provided by the universal law would turn into a chaos of particularity, i.e., preferences and privileges. This way, moral practice if defined by the dialectic of justice and mercy, the eternal tension between two separate sources of ethics – the universal and the particular.

This view might look rather unconventional. However, it seems to be expounded and exemplified in one of the most famous religious and philosophical texts, the *Bhagavad Gītā*. It has been long noted that there are striking similarities between Kantian ethics and *Gītā*, most notably – the insistence on doing one's duty for duty's sake, despite what the desires are inclining the agent to do (Radakrishnan, 1911). Kant establishes this principle in the *Groundwork of the metaphysics of morals* by appealing to reason (GMS; see also Maitra, 2006). For Kant it is the rational, the universal that drives moral decisions, not the particular that is steeped in emotions and considerations of pleasure and pain. Similarly, in *Gītā* Krishna is convincing Arjuna that the only sound reason for action would be duty:

*... Find full reward
Of doing right in right! Let right deeds be
Thy motive, not the fruit which comes with them
(Bhagavad Gita: The Song Celestial, 1965, Chapter 2, p13)^{xi}*

The most clear parallel can be drawn between Kant's ethics of pure reason and *Gītā*'s doctrine of *nīṣkāmakarma*, or desireless action (Matilal, 2002; for a thorough critical review, see Chatterjea, 2002). *Nīṣkāmakarma* appears in the second chapter of *Gītā* (verse 47) as action that is not driven by the agent's

desire of the fruits it is supposed to bring, and continues to be clarified throughout the text as a way of behavior that is not driven by any sort of emotional consideration but performed only because it is the right thing to do. This pertains both to specific contexts, e.g., devotional (*Gītā* XVII/11-13, 17-19), and the general context of acting out of duty (*Gītā*, II/38; III/7; V/11, 13; XVIII/7-9, and elsewhere).

If this were all what *Bhagavad Gītā* could say about moral principles, it could have been both significantly shorter and much more vulnerable to the critique frequently voiced against Kantian ethics, critique that accuses it of the ability to accommodate any rule-compliant principle, as obviously immoral as it can be. *Gītā*'s approach, however, is more complex and sophisticated. A number of philosophers noted the differences between Kant's and *Gītā*'s approaches and explained them by the differences in the ways Kant and *Gītā*'s authors conceive of freedom (Radakrishnan, 1911), by the allegedly different motivations behind the two systems (Maitra, 2006), or by *Gītā*'s scope that suggests "reorientation of the whole life" (Chatterjea, 2002, p141). Yet it seems that the real difference between the two is deeper and more illuminating. Kant argues that the reason must be the sole motive of morality – that much is evident from the *Groundwork*; even though this rigor seems to be significantly relaxed in the later *Metaphysics of morals* (Kant, 1797/1996), where Kant, to the heart's content of many modern virtue theorists, argues for a number of positive duties, yet still insists – ultimately unsuccessfully – on deriving them from pure reason alone. *Bhagavad Gītā*, on the other hand, offers a whole spectrum of motivations: the adherence to the calling of one's social station (Chapter II), behaving in accord with the order of the world (Chapter III), knowledge (Chapter IV), faith (Chapters VII-X and XII), etc. It emphasizes that multiple roads, i.e., multiple motivating factors and variants of moral behavior, can be followed in order to achieve salvation (Chapters XIII and XVII). Yet none of this contradicts the notion of duty. Duty, both in *Bhagavad Gītā* and in Kant, is formal. It refers to the *order of reason* for action, which is *nīṣkāmakarma*, the way of acting *from* duty rather than from desire – not to the particular behaviors or behavioral principles (maxims). As such, it lacks specific content. Without having a maxim to be submitted before the judgment of the categorical imperative, the latter will lead to no action. *Bhagavad Gītā* recognizes this need to provide for the contents that must fill the form of *nīṣkāmakarma* and offers the concept of *svadharma* that has been mentioned above, the concept of specific obligations that come out of our nature, both as members of the human species and as particular people, and its implacement^{xii} in social circumstances. Yet these obligations and the motivations they supply are to be carried out in the *nīṣkāma* way, in accordance with the universal form of acting from duty; it is this way that makes actions moral (cf. Chatterjea, 2002, p135), precisely as in Kant's universalist ethics. This might serve as an illuminating example of resolving the seeming contradiction between the universal and the particular components of morality.

Bibliography

Kant's works – German original:

Kant, Immanuel: Gesammelte Schriften Hrsg.: Bd. 1-22 Preussische Akademie der Wissenschaften, Bd. 23 Deutsche Akademie der Wissenschaften zu Berlin, ab Bd. 24 Akademie der Wissenschaften zu Göttingen. Berlin 1900ff.

Kant's works – English translation:

Kant, Immanuel (1781, 1787/1998). *Critique of Pure Reason*. Cambridge: Cambridge University Press

Kant, Immanuel (misc/1996). *Practical Philosophy*. Cambridge: Cambridge University Press

Sigla:

GMS – Groundwork for the metaphysics of morals

KpV – Critique of practical reason

KrV – Critique of pure reason

MS – Metaphysics of morals

Bhagavad Gītā:

Gītā – *Bhagavad Gītā* in Radhakrishnan & Moore, 1957, pp102-163.

Gītā II/35-36 – Chapter II, verses 35-36

Arendt, H. (2006). *Eichmann in Jerusalem : a report on the banality of evil*. New York, N.Y.: Penguin Books.

Bhagavad Gita: The Song Celestial. (1965). (E. Arnold, Trans.). New York: The Heritage Press.

Chatterjea, T. (2002). *Knowledge and freedom in Indian philosophy*. Lanham, Md.: Lexington Books.

Cialdini, R. B. (2009). *Influence : science and practice* (5th ed.). Boston: Pearson Education.

Cottingham, J. (1978). 'A Brute to the Brutes?': Descartes' Treatment of Animals. *Philosophy*, 53(206), 551-559.

Descartes, R. (1989). *The passions of the soul* (S. Voss, Trans.). Indianapolis: Hackett Pub. Co.

Dostoevsky, F. (1880/2007). *The Karamazov Brothers* (C. Garnet, Trans.). Hertfordshire: Wordsworth.

Floridi, L. (2010). *Information : a very short introduction*. New York, NY: Oxford University Press.

Gilligan, C. (1993). *In a different voice : psychological theory and women's development*. Cambridge, Mass.: Harvard University Press.

- Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science*, 316(5827), 998-1002.
- Hegel, G. W. F. (1807/1977). *The Phenomenology of Spirit*
- Herman, B. (1993). *The practice of moral judgment*. Cambridge, Mass.: Harvard University Press.
- Iacoboni, M. (2005). Understanding Others: Imitation, Language, Empathy *Mechanisms of Imitation and Imitation in Animals, vol. 1 of Perspectives on Imitation: From Mirror Neurons to Memes B2 - Mechanisms of Imitation and Imitation in Animals, vol. 1 of Perspectives on Imitation: From Mirror Neurons to Memes* (pp. 77-100). Cambridge, MA: MIT Press.
- Kant, I. (1797/1996). *The metaphysics of morals*. New York: Cambridge University Press.
- Maitra, K. (2006). Comparing the Bhagavad-Gita and Kant: A lesson in comparative philosophy. [Article]. *Philosophy in the Contemporary World*, 13(1), 63-67.
- Matilal, B. K. (2002). *Ethics and epics : the collected essays of Bimal Krishna Matilal* (Vol. II). New Delhi Oxford: Oxford University Press.
- Milgram, S. (1963). Behavioral Study of obedience. *The Journal of Abnormal and Social Psychology*, 67(4), 371-378. doi: 10.1037/h0040525
- Molnar-Szakacs, I., Wu, A. D., Robles, F. J., & Iacoboni, M. (2007). Do you see what I mean? Corticospinal excitability during observation of culture-specific gestures. *PLoS One*, 2(7), e626. doi: 10.1371/journal.pone.0000626
- Nietzsche, F. W. (1917). *Thus spake Zarathustra* (T. Common, Trans.). New York,: Random House.
- Nishida, K. (2012). Basho (J. W. M. Krummel & S. Nagatomo, Trans.) *Place and dialectic : two essays* (pp. 49-102). New York: Oxford University Press.
- Olson, G. (2008). We Empathize, Therefore We Are: Toward a Moral Neuropolitics Retrieved March 31, 2012, 2012, from <http://neuropolitics.org/defaultnov08.asp>
- Plato. (1997). *Complete works* (J. M. Cooper & D. S. Hutchinson, Trans.). Indianapolis, Ind.: Hackett Pub.
- Radakrishnan, S. (1911). The Ethics of the Bhagavadgita and Kant. *International Journal of Ethics*, 21(4), 465-475.
- Radhakrishnan, S., & Moore, C. A. (1957). *A source book in Indian philosophy*. Princeton, N.J.,: Princeton University Press.
- Rand, A. (1992). *Atlas shrugged*. New York, N.Y., U.S.A.: Dutton.
- Rest, J. R. (1999). *Postconventional Moral Thinking : A Neo-Kohlbergian Approach*: L. Erlbaum Associates.
- Schmitz-Berning, C. (1998). *Vokabular des Nationalsozialismus*. Berlin ;: New York : W. de Gruyter.
- Schulte-Rüther, M., Markowitsch, H. J., Fink, G. R., & Piefke, M. (2007). Mirror Neuron and Theory of Mind Mechanisms Involved in Face-to-Face Interactions: A Functional Magnetic Resonance Imaging Approach to Empathy. [Article]. *Journal of Cognitive Neuroscience*, 19(8), 1354-1372.
- Sripada, C. S. (2008). Nativism and Moral Psychology: Three Models of the Innate Structure That Shapes the Content of Moral Norms. In W. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 1, pp. 319-343). Cambridge: MIT Press.
- Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Trans. Translation newly rev. and edited / ed.). Cambridge, Mass.: MIT Press.
- Wittgenstein, L. (1922/2003). *Tractatus logico-philosophicus*. New York: Barnes & Noble, Inc.
- Wittgenstein, L. (2009). *Philosophical investigations* (G. E. M. Anscombe, P. M. S. Hacker & J. Schulte, Trans. Rev. 4th ed.). Chichester, West Sussex, U.K. ; Malden, MA: Wiley-Blackwell.
- Woody, J. M. (1998). *Freedom's embrace*. University Park, Pa.: Pennsylvania State University Press.

ⁱ See, for example, Haidt, 2007. While Haidt claims that certain population groups have more than two components of morality, he finds justice and mercy to be common across all his subjects.

ⁱⁱ Academic Edition (Akademische Aufgabe), volume 6, page 385. For the *Critique of Pure Reason* (Kritik der reinen Vernunft), pages in the first (A) and the second (B) editions are provided, as it is customary in quoting this book

ⁱⁱⁱ Kant has three more formulations of the categorical imperative, yet this one is the main one, to which he suggests to refer in cases of doubt.

^{iv} This example can also be seen as foundational for the *Metaphysics of Morals* and Kant's ethics of virtue in general.

^v Even though in that case, as it was noted earlier, it will cease to be an expression of mercy

^{vi} The difference between *general* and *universal* rules is principal, even though it is beyond the scope of the current paper. A general rule can be applied by a moral subject multiple times, and he can even make it the guiding maxim of his life. Yet he knows full well that this rule cannot be a universal law of nature. As an example, we can look at the principle of lying to maximize personal gain. An individual can make this the guiding principle of his life, and thus a *general* principle. Yet he understands that this principle cannot be made into a universal law of nature without contradiction.

^{vii} Polytheistic deities can greatly exceed people in terms of their physical abilities, yet they are limited by their physicality nevertheless. It would be interesting to try and understand the connection of this aspect of polytheism with its obvious prevalence during most of the human history in societies with a high degree of religious observance.

^{viii} For the review of the types of information and its processing see Floridi, 2010.

^{ix} For the discussion on spontaneous and scientific concepts see Vygotsky, 1986.

^x The opinion that animals are merely automata is attributed to Descartes himself, following his reasoning in *Passions of the soul* (Descartes, 1989). The correctness of this interpretation of Descartes' views has been contested, however – see, for example, Cottingham, 1978.

^{xi} For a more scholarly, even though less poetic translation see Radhakrishnan & Moore, 1957, Chapter 2, 47 / p110.

^{xii} Possible parallels with Nishida Kitaro's *basho* concept can be seen here; see Nishida, 2012