

## Why Common Sense Morality is Not Collectively Self-Defeating \*

Piotr Boltuć  
*University of Illinois at Springfield*

**Abstract.** The so-called Common Sense Morality (C) is any moral theory that allows, or requires, an agent to accept special, non-instrumental reasons to give advantage to certain other persons, usually the agent's friends or kin, over the interests of others. Opponents charge C with violating the requirement of impartiality defined as independence on positional characteristics of moral agents and moral patients. Advocates of C claim that C is impartial, but only in a positional manner in which every moral agent would acquire the same relational characteristics if that agent was in a certain relationship to the given moral patient. The opponents of C reply that a theory that allows for positional characteristics is self-defeating; it violates the requirement of prescriptivity due to its inability to provide moral recommendations *what should happen all things considered*. Advocates of C retort that a moral theory should be prescriptive by telling every agent what to do, not what should the joint outcome of those activities be. In this paper I analyze the last two moves of this debate: the objection that C is self-defeating and the reply that there is a plausible moral theory (C) that accommodates positional characteristics of special moral reasons. I argue that the last move wins. In the process I sketch out a theory able to accommodate agent-relative moral reason.

### 1. Introduction

In order to accommodate moral reasons generated by friendship, love or kinship a good moral theory, it has been claimed, *should* incorporate non-instrumental agent-centered moral reasons. In this paper I defend a weaker claim, that a moral theory may include special moral reasons.

I take this argument to be important for the following reasons: While many authors argue that special moral reasons, such as the reasons of friendship, should be a part of a good moral theory (among others: Cottingham, 1986, 1988; Scheffler, 1997, 2001; Blum, 1988; Nagel, 1986, 1991; Parfit, 1991; in a moderate way Friedman, 1991, 1993; Archard, 1995), and some present a theoretical framework for

---

\* I am grateful to Charles Byrne, R.G. Frey, Ellen Frankel Paul, Ned McClennen, late Michael Robins and my students at the University of Illinois at Springfield, especially Melissa Winkel, for their comments on various versions of this paper.

such theory (Sen, 1982, 1983, 1993, 2000; Portsmore, 2003, 2007), the standard objections to this position presented by Derek Parfit (1979, 1984; also Regan, 1983) almost three decades ago still await satisfactory responses. If those objections can not be addressed, the whole project of ethics of special reasons becomes dubious.

### 1.1. PARFIT'S CHALLENGE

Ethics that allows for non-instrumental moral reasons is often called Common Sense Morality (C) (Sidgwick, 1874; Parfit, 1984). Parfit accuses C of collapsing into the Present Aim Theory (P), a theory that requires agents to attend to one's present aims only (and thereby into Full Relativity). He also makes a separate claim that C is indirectly collectively self-defeating so that it leads to inconsistent moral recommendations. In this paper I argue that C is not inconsistent. I have argued elsewhere that it is plausible (Boltuc, 2001; 2005).

Parfit's rejection of C relies on two arguments.

- (a) The argument presumes that only two kinds of moral theory are consistent: Impartial Morality (M) and Present Aim Theory (P); P is implausible; Common Sense Morality (C) collapses into P; therefore, C is implausible. This leaves M.

In responding to the above argument I shall distinguish between impartial (M-style) and particular (P-style) theories, which distinction belongs to moral motivation (or action theory), and the content of moral beliefs one acts upon. I shall claim that such content, whether motivated by M-style or P-style reasons, may also include mixed partial and impartial considerations the way C does. The notions of universal benevolence and self-interest, which create the dichotomy explored by Parfit, shall be viewed as boundary conditions of the notion of mixed benevolence. Such benevolence extends to every human (or sentient) being, but it is non-homogeneous—this means that moral benevolence may be stronger toward those closer to us in a morally relevant sense. Ethics based on this notion of benevolence constitutes a version of C.

Parfit's other argument relies on the following point:

- (b) C leads to mutually inconsistent recommendations that it gives to different agents; thus, C is incoherent.

In this argument Parfit claims that systems that allow for agent-centered moral reasons are unable to tell us what should actually happen if two or more agents have conflicting moral reasons in a given situation. I present Amartya Sen's reply to this argument that consists of two

intertwined points: that the ordering of preferences does not need to be complete and that agent-relative reasons make such ordering impossible. Sen bites the bullet and presents a moral theory that stops short of trying to tell moral agents ‘what should happen’ and instead provides moral rules of, sometimes competitive, interactions among agents.

My answer to (b) follows Amartya Sen. My own contribution is limited to the answer to (a) and developing the consequences of Sen’s answer to (b) for competitive games.

## 2. Present-Aim Theory

In criticizing argument (a) I demonstrate that it depends on mixing present desires about any past, present, or future desires (P) with present desires the content of which it is to take into account solely present desires (PP). If we distinguish P and PP properly Parfit’s argument against C fails.

Parfit begins by asserting that the Self-interest Theory (S), better known as moral egoism, is not directly individually self-defeating; this is because it does not claim to be impartially concerned about everybody. Parfit attempts to convict S on a lesser charge. He accuses S of being unable to jointly defeat Impartial Morality (M) and the Present Aim Theory (P). S is able to defeat each of them but not both—since S lies between M and P, the arguments that work against one of them also support the other.

In the second move of his argument, which makes it relevant for this paper, Parfit argues that C is indeed a version of S; hence, the charges launched against S also reach C.

What is the Present Aim Theory? Parfit defines it as a theory that “tells each to do what will best achieve his present aim” (Parfit, 1984, pp. 92, 117). The difference between P and S is due to the fact that I may either do what best achieves my present aims (following P) or “do what will best achieve [...] my aims over time” (following S). If I choose the first option, “over time I may then do worse” in terms of P (applied at that future time) than if I had followed S. Hence, according to Parfit, P is inter-temporally self-defeating. Yet, Parfit claims, advocates of S, or any theory which assumes Reason Relativity (indexicality), cannot build an argument against P. This is because P does not claim to be inter-temporally successful (hence, it does not fail on its own terms).

Parfit emphasizes that this defense is structurally just like a defense given by S against the charges that it is inter-personally self-defeating. S is not committed to utility maximization among different agents, just

like P is not committed to utility maximization over time. Therefore, the argument continues, “if it is a good reply that S does not claim to be collectively successful, why can the Present Aim Theorist not make a similar reply?” (Parfit, 1984, p. 94).

Parfit rejects the potential reply that there are differences between impartiality among different persons and between different time-slices of the same agent. Time-relativity and person-relativity stand or fall together. If they stand, the new theory leads to Full Relativity; this vindicates P since it involves both time-and person-relativity. If they fall, the new theory is compatible with impartial morality (M). Parfit concludes, “the only tenable theories are morality and Present Aim Theory, for only they give to ‘I’ and ‘now’ the same treatment” (Parfit, 1984, p. 148).

### 2.1. DESIRES ABOUT DESIRES

There is a problem with Parfit’s argument. He considers P in its radical version. Radical P is a conception of the good in which solely the agent’s present desires count.<sup>1</sup>

However, the list of theories of the temporal locations of desires considered by Parfit is too short. We may have a theory of the good F where only the agent’s future desires count. We may also have a conception K in which only the agent’s past desires count. Few people hold F or K, but these theories are not vastly implausible; also few people, after all, hold P (or S) in their pure form.

What is the relationship among P, F, and K? F can be seen as a theory in which an agent’s present desire P is such that she cares *now* solely about all of her future desires, and she cares about them equally. K can be seen as a theory in which a present desire of a moral spectator is such that she cares *now* only about all her past desires equally (or she cares about desires that occurred before point t in the past). Most interestingly, we may also have a theory PP *in which a present desire of a moral spectator is such that she cares now solely about all her present desires.*

Parfit misses the point that present desires may have temporality fixed as a part of the content. Consequently, Parfit conflates theories P and PP. P is a theory telling us to follow whatever defines our present aims. F, K and PP define a temporal aspect of the content of those aims. Hence, according to P, a present desire can be to realize present first order desires (which is the only option satisfied by PP) but it can

---

<sup>1</sup> Parfit tries to demonstrate that counting an agent’s past desires leads to paradoxes, but we may have a theory S in which only present and future desires count.

also be to realize one's past desires (described by K), or future desires (described by F).

If this is the case, then we can have hybrid theories of the content of present desires in P. For instance, Władysław Tatarkiewicz's present aim at time  $t$ , when he was writing the book *On Happiness*, could have been to satisfy present, future and past desires in some elaborate proportion, *which is actually discussed in that book* (Tatarkiewicz, 2007; 1976). Let us call such hybrid theories Tatarkiewicz Teleology (T). Hence, we may have a continuum of temporal preferences where F, K and PP are just the extremes; whatever lies in between is a T-type theory.

Now we may come back to Self Interest (S) and characterize it as a certain mix of PP and F. This is because self-interested people often care about both some of their present and future desires; hence, S is a version of T.

## 2.2. P AND M THEORIES VERSUS THEIR CONTENT

If it is true that Parfit neglects the fact that present desires may have temporality fixed as a part of the content, it is also true that person-dependent moral theories may have agent-relativity, or agent-neutrality, fixed as a part of their content. As Parfit claims, *time-relativity and person-relativity stand or fall together*. If I am an agent motivated just by my present desires, it is conceivable that the content of my present desires is to bring peace and happiness to the world, or some other desire of the M-kind. This is a PM theory such that Present Aim-type motivation is indeed aimed at realizing present moral desires of the M-kind. We can also have a PT theory, where, for instance, somebody has a present personal desire (which she does not try to generalize) with the content such that, for instance, every parent should attend to her children. Finally, we can have a PP theory, where the content of one's present desires is to do whatever one's present self-interest, or ongoing desires, may tell her, which is a PP theory. It seems that PP is really what Parfit expects P to be, but there is no reason to think that this theory is privileged over PT or PM.

Parfit seems right that the sole alternative to P (or rather to Full Relativity implied by it) is M. Yet, this is the case at the meta-level of moral motivation, not at the level of the content of moral prescripts, or of moral desires, recommended by P or M. Hence, by analogy with PM, PP and PT, we can also have MM, MP and MT; in these instances the content of a moral theory, accepted on an impartial basis, is either M, P or T. A person acting from the impartial viewpoint may come up with an objective, impartial moral justification for a certain content,

while that content is not identical with M, but with P (the so-called *objective moral egoism*, Hampton, 1998) or T (most advocates of moral partiality argue that C is required by impartial ethics so that C is an objective moral requirement, for instance, that every friend should take care of her friends to a larger degree than of other people).

This observation, which seems central to my argument, shows that, even if we begin the argument with a sharp distinction between the theories that are strictly homogenous, such as Impartial Morality M, and those that focus only on the subject (P) we can factor in various preferences in the content of those present desires. In fact when the content of desires in P is identical with M we end up with a meta-theoretical Present Aim Theory whose aim, however, it is to act impartially (to follow M). This option is not empty since, in action theory, meta-theory P could provide the motivational force to realize M.

We need to develop one point mentioned above concerning T-type theories in ethics. In general, we can have exactly two types of such theories.<sup>2</sup> First, we may have a P-type theory, which motivates agents by present interests, where it is subjectively desired that we care only about the requirements of Common Sense Morality (C); this is a PT theory. Second, we can have an M-type theory, which motivates agents by impartial considerations (M) whose content is also T-style; this is an MT theory claiming that everybody should act according to C. In both cases we attain the Common Sense Morality (C) but within a different meta-theoretical framework (subjective in PT and objective in MT).

Let me unpack this heavily meta-ethical section. Parfit's dichotomy between P and M: either impartial (M) or agent-centered (P) justification of morality (or motivation to follow its prescripts) plays a role in the theory of moral motivation. We can have two basic kinds of moral motivation. The first one, related to M, that tries to view every agent impartially, *sub specie aeternitatis*, from the viewpoint of Hare's *Archangel* (the motivation for such theory is some kind of a *moral sense* or a desire to bring about impartial justice). The second one, related to P, is ego-centered in terms of motivations and follows Hobbes, Gauthier and Hampton in trying to view moral actions as an attempt at coordination among various self-interested agents. Such issue of motivation, of either kind, is often viewed today as a part of action theory. Yet, we need to emphasize, that, at the level of the content of moral theories (with their motivational component being of either of the above two kinds), one can choose among various structures: not only impartial

---

<sup>2</sup> We could also have a second-person meta-ethics, which seems to be proposed by Buber, Lévinas and some of their followers, but this is a rare and controversial option (Boltuc, 2008).

(M-style) or self-centered (P-style) but also various mixes of the two (T-style).

### 2.3. TIME- AND PERSON-RELATIVITY

Since according to Parfit time-relativity and person-relativity in ethics stand or fall together, we can unify the above arguments for both temporal and moral aspects. Below I present the moral aspect since it is relevant for the rest of this paper.

As shown in the previous section, we can have three kinds of theories based on P: PP, PM and PT. The PP theory designates P motivation to accept P content; it tells us that we are (or, should be) motivated by our present desires, where the desire is to do what one presently feels like doing. PM theory also tells us to be motivated by our present desires; yet, here the desire is to treat each of one's own past, present and future desire equally (or even to treat all desires of all people equally, which is the basis of a moral motivation to treat everybody impartially advocated by Parfit).<sup>3</sup> Finally, we have PT theories, according to which again we are, or should be, motivated by our present desires; yet now the content of those present desires is (or should be) to follow a somehow structured mix of our present and other desires.

We can also have three kinds of theories based on M: MP, MM, and MT. The first is a theory of objective egoism that claims that, from the objective viewpoint, everybody should follow one's present desires. The second one, MM is the traditional impartial ethics which requires, from an impartial viewpoint, for everybody to treat moral agents equally. Finally, MT is the theory that recommends impartially for everybody to attend to some special desires (or duties). In the area of ethics MT designates, among other T-type theories, Common Sense Morality (C) the way Marilyn Friedman does (Friedman, 1991; 1993). T-type theories, including C, are immune to Parfit's objection and therefore they do not fall into P theories.

---

<sup>3</sup> Would anybody have such a desire? Utilitarians often claim that one should be impartial towards present and future desires (Sidgwick, 1974), while there are decision-theoretic (Elster, 1979; McClennen, 1990) and, different, conservative (Kirk, 1985) reasons to be impartial toward one's past desires. Here is the decision-theoretic reason in a nutshell: If I know that nobody cares about their past desires, or about the desires of dead people, then I must build elaborate structures, such as trust funds, in order to enforce my present desires in the future (when they will become past desires); this increases decision-theoretic enforcement costs (Elster). M, if applied to ethics, not just to temporal preferences, agrees with some schools of virtue ethics which claim that a morally good person is the one whose present desire is always identical with M; they put M at the level of present desires not metaethical, motivations so that theirs are PM theories.

If this is true, Parfit's argument aimed at showing that C collapses into P is unsuccessful. This is because P and M can be viewed as *tertium non datur* alternatives only at the meta-ethical (or meta-temporal) level. At the level of the content of moral reasons we can have P-style, M-style and also T-style theories and the latter kind of theories (PT and especially MT) accommodate C.

### 3. Inconsistency of Recommendation

In part I of the paper I responded to Parfit's claim that Common Sense Morality (C) collapses into Present Aim Theory (P). Now, I address his claim that C leads to mutually inconsistent recommendations it gives to different agents. In the process I sketch out a broader framework of moral theory able to accommodate special moral reasons, as well as the general ones. Difficulty in formulating such a framework seems to be the principle remaining objection against C.

I take the following path: First, I present Parfit's inconsistency of recommendation argument. Next, I present Sen's three kinds of agent-relativity and their binary relationship; this is to show a structure of moral theory that is an alternative to M. I also discuss some reasons for adopting such a structure of moral theory, although the structure does not allow for a complete ordering of outcomes. Later, I argue that there is no reason to require a complete ordering of moral options (or of outcomes of moral actions) since systems that involve such ordering are unable to account for competition. I also argue that competitive situations are a vital subject of moral theory. Finally, I draw the structure of consequence-based moral theory that is compatible with C.

#### 3.1. PARFIT'S ARGUMENT FROM INCONSISTENCY OF RECOMMENDATION

Parfit's *inconsistency of recommendation* objection is based on the observation that the system of particular moral obligations (C) assigns different aims to different moral agents; hence, in the long run C leads to conflicting claims. Thus C is indirectly collectively self-defeating. This objection to C was earlier posed by Sidgwick (1874, pp. 337–61; Parfit, 1984, pp. 95–112). C fails since the world where people try to achieve incompatible moral ends (and so, need to compete in order to achieve them) is clearly sub-optimal to the world where the results are assigned by impartial ethics (M), or so the argument claims.

Parfit introduces a few distinctions that help define the exact manner in which theories can be self-defeating. A theory (T) can be indirectly



or directly self-defeating. The former refers to an attempt to achieve the aims given by T, whereas the latter successfully obeying T. A theory is directly self-defeating if it gives “different people different aims” (Parfit, 1979, pp. 534–35). Parfit also distinguishes between theories being directly *individually* self-defeating and directly *collectively* self-defeating, depending whether the conflict occurs among inconsistent goals of a particular agent or various persons bumping into each other in an attempt to pursue their respective goals.

Parfit points out that, if two agents interact while having different aims, we have four options that fit with the classic decision-theoretic scheme of the Prisoner’s Dilemma: either I cooperate and the other agent cooperates; I cooperate while the other agent does not; she cooperates and I do not; both of us fail to cooperate. Parfit argues that self-interested agents fall into the trap of Prisoner’s Dilemma, so that it is individually rational for them to choose the *lower equilibrium*,<sup>4</sup> which is a paradigm of theories where acting on individual rationality is self-defeating. (This is because *third person* rationality would have resulted in an outcome better for each player but the players have no reasonable way to attain it within the framework of individual rationality, except for long reiterative games, McClennen, 1990). Hence, for instance, egoism is self-defeating.

Parfit hopes for a moral theory to do better; it must assign mutually consistent goals to all moral agents. Otherwise, while competing for realization of their respective goals moral agents would frustrate those attempts by others. Even if the only goal of moral agents was to achieve their individual aims (defined by egoism), it would be more rational for them to cooperate and therefore to have an impartial set of mutually consistent goals. This paradox is true not only for self-interested, egoistic agents but even for what Parfit calls ‘self-referential altruism’ (which seems to be the prototype of Sen’s idea of moral agents acting on agent-relative moral reasons). Parfit demonstrates that it causes the same paradoxes (Prisoner’s Dilemmas) as egoism. He argues that the structure of the Prisoner’s Dilemma replicates both for agent-neutral and agent-relative aims if moral theory fails to assign the same goals to all agents (Parfit, 1979, p. 543).<sup>5</sup>

---

<sup>4</sup> The following discussion presumes basic understanding of the Prisoner’s Dilemma. This famous paradox of collective rationality is a part of any introduction to decision theory.

<sup>5</sup> In this area Parfit’s work started a major discussion pertaining to Prisoner Dilemmas and their self-defeating nature that started with Ian Elster’s work (1979; 2002). Edward McClennen (McClennen, 1990) arguing that rational agents can aim at the high equilibrium counter through *resolute choice* to individual rationality narrowly understood. They must be able to restrain themselves from self interested

Parfit argues that moral theory must make consistent recommendations concerning the outcomes, not just the obligations, of various agents. For him, a theory (P) that assigns agent-relative goals “is not a moral theory.” Parfit’s criticism of all agent-relative theories pertains not only to ‘common sense morality’ but also to any (moral) theory based on self-interest such as most forms of contractarianism, which assume that agents are self-interested.<sup>6</sup>

To sum up, the main criticism that Parfit directs toward C is that it produces the following situation: (1) it gives different agents different aims, (2) the achievement of each person’s aims partly depends on what other persons do, yet (3) “what each does will not affect what these others do” (Parfit, 1984, p. 95).<sup>7</sup> Parfit’s general point is that special obligations (e.g., friendship), which require us to give priority to helping our relations over strangers, lead to collisions among different moral agents since they require various agents to favor different individuals and their interests (Parfit, 1985, pp. 443–44).

### 3.2. SKETCHING OUT MY REPLY

Whether Parfit’s *inconsistency of recommendation* argument holds depends on the assumptions that all evaluations can be compared, which requires a complete ordering of states of affairs. But there are objective reasons (ascertainable at meta-M impartially) to view certain situations from the viewpoint of an agent, in terms of Sen’s *doer-relativity*, *viewer-relativity* and *self-evaluation relativity* (Sen, 1982; 1983); the gist of those ideas will be presented later in this paper. In such cases there is no acceptable third-person objective perspective (the second ‘M’ in the MM impartial ethics) to refer to at the content level and aggregation of utilities is not always possible, or morally desirable. This entails the rejection of the standard utilitarian model of moral theory.

The model proposed to replace it is broadly rule utilitarian but includes agent-relative evaluations; it follows the framework of what Sen calls ‘consequence based moral evaluation’ (1983; 2000), or broad

---

myopic actions if their long-term self-interested rationality requires that they cooperate. McClennen’s results may have been confirmed by Jason Alexander’s computer simulations (Alexander, 2003) in the framework of evolutionary game theory but only for reiterative games (interactions that come in long sequences with the same agents involved).

<sup>6</sup> Scanlon (2000) is an exception to this rule (see p. 8).

<sup>7</sup> Parfit’s third point, in its present formulation, covers solely cases of coordination games, whereas he should have also covered cases of direct competition which are clearly allowed by conditions (1) and (2). In these cases, “what each agent does” will affect what others do, but in a negative way. Parfit ignores zero-sum games, which is vital for my further discussion.

consequentialism (1993). The model, which is a version of C, allows morally for competition of interests, which Sen views as a major cause of its better fit with reality than M. In response to Parfit's argument against C based on inconsistency of recommendation I shall develop the following argument, which is based on the two mutually related objections sketched out above: Since some moral values are agent-relative, and C, but not M, is able to provide correct theoretical account of this fact, while agent-relative reasons do not allow for a complete ordering of outcomes, there are strong reasons to doubt that we should always have such ordering.

In order to make this argument we need to define agent-relativity and its role in moral theory. Since in this paper I argue that a moral theory *may* include special moral reasons, not that it has to, it is sufficient for us to argue that a moral theory *may* involve agent-relative reasons.

### 3.3. THREE KINDS OF AGENT-RELATIVITY

In my view, Sen's version of an agent-relative, outcome-based morality is the principal moral theory worked out in some detail that provides support for C. It seems to rely on T. Nagel's theory of first- and third-person perspective but is presented with more analytical rigor. This allows Sen, best known as a Nobel Laureate in Economics, to respond to Parfit's objections to C more in detail than Nagel (1986; see also McNaughton, 1988).

Sen emphasizes that observations, beliefs and actions can be objective but *viewpoint (position) dependent*. They involve "the parametric dependence of observation and inference on the position of the observer." He distinguished three aspects of agent-relativity: *doer relativity*, *viewer relativity* and *self-evaluation* relativity. In his earlier papers (1982, 1983) Sen does not commit himself to accepting agent-relativity, but rather considers, with a sympathetic eye, what formal consequences follow from arguments for agent-relativity "if they are accepted" but he commits himself to this view later on (2000). According to Sen, although agent-neutrality has been a feature of consequentialism because of historic contingencies, there is nothing in the very structure of consequence-based moral theories that would require agent-neutrality.<sup>8</sup>

Sen makes it clear that what philosophers refer to as agent-relative values actually constitutes several categories with different formal char-

---

<sup>8</sup> Jonathan Dancy (1993) seems mistaken, in his otherwise excellent work, in believing that agent-relative consequentialism cannot be achieved.

acteristics. He distinguishes three types of agent-relativity, defined as negations of the following neutrality claims (Sen, 1982, p. 21):

(DN) *Doer neutrality*. Person  $i$  may do this act if and only if person  $i$  may permit person  $j$  to do this act,  $A_i(i) \Leftrightarrow A_i(j)$ .

(VN) *Viewer neutrality*. Person  $i$  may do this act if and only if person  $j$  may permit person  $i$  to do this act,  $A_i(i) \Leftrightarrow A_j(i)$ .

(SN) *Self-evaluation neutrality*. Person  $i$  may do this act if and only if person  $j$  may do this act,  $A_i(i) \Leftrightarrow A_j(j)$ .

Sen has discovered an interesting formal regularity. These three kinds of agent-neutrality are bilaterally dependent on each other, which means that any one form of agent-relativity entails one other form: “[I]f any one type of agent-relativity is satisfied, at least one other type of relativity will also obtain” (Sen, 1982, p. 22). Hence, if we discover the existence of one of these kinds of agent-relativity, we may presume the presence of at least one other kind.

### 3.4. REASONS FOR AGENT-RELATIVITY

What are the main reasons to adopt agent-relative values? Sen distinguishes the *deontic reasons* (i.e. “agent-relative reasons for each individual not to maltreat others himself in his dealings with them [...] by violating their rights, breaking his promises to them etc.,” Nagel quoted in: Sen, 1983, p. 120) and *reasons of autonomy*, (based on “desires, projects, commitments and personal ties”). There are two kinds of reasons of autonomy: *reasons of integrity* and *reasons based on personal ties*.

*Reasons of autonomy* include *reasons of integrity*. Those reasons consist of: *integrity respect* (“a general respect for integrity no matter whose”), which is not even agent-relative, and *integrity responsibility* (“personal responsibility for one’s own integrity”) (Sen, 1982, p. 26). The difference between reasons of autonomy and deontic reasons is that, in the case of deontic reasons, I refrain from certain acts because identifiable others have an agent-relative claim on me not to do certain things to them, while in the case of integrity responsibility, I have an agent-relative obligation (to myself) to keep my own integrity intact.

The other kind of reasons of autonomy singled out by Sen are *reasons based on personal ties*. These reasons are also of two kinds: *tie aims* and *tie respect*. Reasons of *tie respect* presume that everybody should value the acts of people who value their special ties (for instance, everybody should value parents who care for their own children). This

is an instance of MT theory sketched in Part I. These reasons are *doer relative*, since these are reasons for parents whose actions toward their own children are valued, while other people could not replace them even if they acted in the same manner. It also involves *self-evaluation relativity* because father A may do a certain thing to his own children but father B may not do this to A's children. However, *tie-respect* does not involve viewer-relativity (it is not an IT theory) since everybody may value parents benefiting their own children.

Reasons of *tie aims* involve different people having different aims; for instance, each person's aim is to benefit her own kin. These reasons are not *doer relative* since respective parents do not care who actually performs the act of benefiting their children; they just want their children to get the benefits. But tie aims are both *viewer relative* and *self-evaluation relative*—they involve evaluator-relativity. They are *viewer-relative* since X may do something to benefit his own children and Y may try to stop him (if X and Y are in a competition). They are also *self-evaluation relative* since agent X may have dominant reasons to evaluate a certain act as good (e. g., that it is a good thing to help X's friend), while the same act may not be good for Y.

Sen also introduces the notion of *aimed tie respect* which involves both *tie respect* and *tie aims*. An example of *aimed tie respect* is favoring one's own children while also maintaining that benefits for them must be produced exclusively by oneself. *Aimed tie respect* stands out as involving every kind of agent-relativity. It inherits *doer relativity* from *tie respect* (the benefits must be produced by the kin), *viewer relativity* from tie aims (I may help my kin even if others do not have reasons to permit this), and *self-evaluation relativity* that comes from both.

### 3.5. EVALUATOR-RELATIVE OUTCOME MORALITY C

Agent-relativity can be accommodated in the general framework of C-compatible broad consequentialism, which is formulated in the following way. Sen defines function  $G_k^i(x)$ , which can be read: "the moral value that in the opinion of person  $k$  (parameter) should be appropriately attached to state  $x$  (variable) by person  $i$  (variable)" (Sen, 1982, p. 33). In a particular case  $k = i$  (observer  $k$  is identical with agent  $i$ ) so that  $G_k^k(x)$ . Evaluator-relative valuation of outcomes is sensitive to differences in the position that different people occupy *vis á vis* states to be evaluated. Since these positions may differ, it may happen that different people evaluate the same state differently.

The point that Parfit objects to is that the same morality recommends different and, on some occasions, incompatible solutions. Yet, in the context of Sen's formulation above, Parfit's accusation of

inconsistency turns out to be incorrect because these various recommendations are coherently interpreted as ‘positional statements.’ By *positional statements* Sen understands statements reflecting the view of the circumstances from the position of the evaluator. Various things are right for people who hold different places in a given situation because, to follow Jonathan Dancy (1993), different moral properties show themselves as salient from different viewpoints. Sen, Dancy, Pargetter and others point out that a moral perspective involves “positional fixity” since it is based on the existential characteristics of an agent (Sen, 1982, pp. 220–21).<sup>9</sup>

Sen shows that people occupy their places not just descriptively but also evaluatively; it is one’s moral feature that she is a friend, a parent, a compatriot and C recognizes those facts while M is unable to do so. It is at least plausible that one’s correct moral viewpoint is not always impartial, so that Hare’s benevolent and omniscient Archangel (the quintessential moral spectator of M (Hare, 1981) is not always the best judge of what a given agent should do. It is sometimes morally wrong to be impartial if to hold such a perspective, one has to violate one’s special (agent-relative) moral reasons such as the reasons of autonomy, including the ties with one’s family, friends, compatriots and other groups of particular others. This last view is no longer controversial in analytic moral theory (Friedman, 1993, Scheffler, 1997; Blum, 1988; Cottingham, 1986, 1988).

### 3.6. MORAL RECOMMENDATION AND INTER-PERSONAL MORAL DILEMMAS

With the background of Sen’s moral framework based on agent-relativity let us come back to Parfit’s *inconsistency of recommendation* argument against C. His objection is similar to Tom Regan’s criticism of Sen’s approach. Regan argues that “divergent points of view require conflict” (Regan, 1983, p. 107).

Replying to this objection, Sen declares that, assuming moral realism, the absence of conflict would not be a merit of a moral theory if we accept moral realism:

we are not discussing the choice of an instrument. If the correct view happens to incorporate inter-positional differences and conflict, then clearly it will be incorrect to insist on absence of conflicts. (1983, p. 126)

Yet, while Sen accepts moral realism, he is agnostic about moral cognitivism, which may need to be reformulated if we adopt “the positional interpretation of moral statements” (p. 118). If such a theory is

<sup>9</sup> In this Sen’s position is similar to the communitarian fixity of one’s viewpoint in social/political philosophy.

supposed to describe moral obligations which are independent of this theory, and the obligations actually conflict, a moral theory should not skirt this conflict.

Particular moral reasons open up the possibility of inter-personal moral conflicts (sometimes termed *inter-personal moral dilemmas*). This is because of possible differences in the general balance of reasons, general and particular, that different moral agents may possess. Let us analyze a pair of examples that will help us highlight this point: Imagine John and Barry who are high school students with the same moral reasons JM and BM. But once they establish their families John acquires a special moral reason (a duty) to help his son, Alex. Hence John's moral reasons are now JMA. At the same time Barry has a daughter, Florence, and his moral reasons are now BMF. If what John owes to Alex, and what Barry owes to Florence conflicts then JMA and BMF (the best moral reasons John and Barry have) may conflict; John and Barry may be morally required to compete to realize their *prima facie* duties of parenthood that they hold to their respective children. For instance, John and Barry may compete for the one spot in a certain kindergarten that Alex or Florence may occupy.

Those special duties may be of the same kind (John and Barry both have the duties of parenthood in this example), but they may differ, which can lead to direct conflict. For instance, to transform the famous Kohlberg case, the *Heinz' Dilemma* (Kohlberg, 1984; Gilligan, 1982; Flannagan and Jackson, 1987; Duska, 1975), John may still be a father of Alex while Barry now is the local sheriff. It may be that Alex needs an expensive medication that John is unable to buy, or to raise the needed money timely, but he may be in a position to steal the medicine. If Alex's ailment is grave enough, and the harm that John would likely cause in the process of appropriating the medication (including the harm to those who would not be able to buy it since it would be stolen) is negligible enough, it may turn out that John will be morally obligated to steal the medicine. At the same time Barry the sheriff has now the special reason L to defend law and order. While it is possible that the general moral reasons of charity that Barry (like everybody) has, tell him, all things considered, to turn a blind eye on John stealing the medicine, or even to assist him, it is also possible, likely even, that the reason to keep law and order dominates and that Barry ought to stop John (this is what sheriffs are for).

According to broad consequentialism, the balance of general harm brought about by John's act of stealing the medicine, versus the balance of harm that comes about if John fails to steal the medicine, including the harm to Alex (who may suffer or die due to the lack of medication), needs to be weighed by every agent, including John

and Barry. Yet, this balance should not be weighed equally. As Sen points out, John has agent-relative reasons (of parenthood) to count Alex' harm for more than Barry, whereas Barry the sheriff has special reasons (based on his professional ethics) to defend law and order more. Those are unequal reasons though none is absolute. If John viewed his reasons of parenthood as absolute he would feel justified in bringing any amount of harm to others in order to save Alex; while bringing about moderate hardships on others may be excusable, and indeed required on him under the circumstances, he is not morally allowed to bring major hardships on anybody. Also, if Barry thought that his duty to defend law and order were absolute, he would prevent any minor transgression whatever the consequences, like a police officer trying to stop a slightly speeding car that he knows is attempting to bring a dying person to the hospital (this *kind of* Kantian attitude would make us report people wanted by the Nazi or Stalinist regime on the basis of pure legalism). Hence, no duty is absolute since there are always limiting conditions beyond which enforcing such duty would be morally wrong. What counts as such limiting conditions depends, in part, on the positional characteristics of each agent (e.g., police officers have a special duty to prevent crimes, and parents another special duty, to help their children).

To take an example from business ethics, it would be morally wrong to advise your company to withdraw from a perfectly competitive bid (a zero-sum game) which it is likely to win, just because the competitor's winning may produce a slightly higher balance of general utility. As an advisor to a particular company, you ought to see the situation in an agent-relative way, primarily from the point of view of your company, since this comes from your allegiance to the company, even if contractual and temporary, although agent-neutral reasons (such as major environmental damage to the broader society) should also be taken into account (Porter, 1998).

This difference in moral reasons, that *all things considered* various agents have, results in inter-personal moral dilemmas in which each agent acts on the best reasons available to her but they clash in real life. Parfit claims that this is a counterintuitive result, but Sen shows two things: that such is the only descriptively acceptable account of moral life, and that such moral theory is consistent (despite its being unable to provide final outcomes). According to such theory moral agents may sometimes be morally required to compete.

Inter-personal moral conflicts require a structure of moral theory (C) that is able to accommodate such conflicts of moral reasons. Hence, it is possible for a moral theory to maintain that general reasons do not always constitute the dominant moral reasons an agent ought to act



upon, although they must be taken into account, especially in cases of acutely harmful outcomes.

### 3.7. NO COMPLETE RANKING OF EVALUATIONS

Agent-relative moral reasons (both deontic moral features, as well as those related to reasons of autonomy and special ties) do not allow for invariance among different agents. Hence, a system C able to accommodate those vital considerations must stop short of such invariance. Utilitarianism confines its considerations to *cumulative outcomes*, and insisting “that all consequences other than utility consequences be ultimately ignored” (Sen, 2000, p. 501), which implies invariance among different agents. The exclusion from moral evaluation of the aspects of the state of affairs other than utility is based on an assumption that such aspects are not morally valuable but this is unwarranted. Sen shows that we can provide agent-relative utility to various deontic and agent-based constraints, such as broken promises, friendships and other special morally relevant ties.

Sen’s *consequence-based system of moral evaluation* has two specific features. One of them, agent-relativity, involves the lack of a complete inter-personal ranking of preferences; the other feature, the mix of deontic and traditional consequentialist reasons, may not involve that. Let me present the latter feature first in order to better distinguish it from the former, agent-relative, aspect.

Sen advocates the mixed model of moral evaluation that takes seriously the fact that a moral agent needs to take responsibility for one’s choice and that the value of such choice is dependent on various positional ramifications (such as personal relationships to the persons involved). This objective involves both deontic and consequentialist aspects, since one’s choice has both; it is deontic in terms of one’s intentions and consequentialist in terms of outcomes they produce. Sen claims that outcomes must be defined, in part, through actions they result from (Sen, 2000; Portmore, 2003). In order to achieve this, consequences must be understood broadly so as to include traditionally deontic features, now treated as objects of evaluation, such as *betrayals or satisfied duties*. The state of affairs with the same amount of pleasure may still have less utility, Sen-style, if it involves negatively valued *deontic objects*, such as broken promises than the state of affairs in which the same non-deontic utility (in this case disutility) comes about. Such states of affairs can be accounted for by M since they treat such products of deontic acts as just some of the objects evaluated within standard consequentialist calculus. Hence, this broadening of the scope

of traditional utilitarian consequentialism has no direct bearing on the debate between C and M.

While the first striking feature of Sen's proposal is that it combines deontic (or integrity-based) features with the traditional utilitarian considerations in a broadly consequentialist framework, the second feature of the system, the agent-relative character of many deontic and integrity-based evaluations, does not allow for invariance among different agents. This last feature is based on the fact that people have particular, agent-relative, as well as general moral reasons to act. As we remember from the John and Barry case, it seems morally wrong, at least in some cases, for John not to promote his child. This is true not only for dramatic cases, such as *Heinz Dilemma*, but primarily for the ethics of everyday life. It is morally wrong for John not to acknowledge any special reasons to want his kid to win the pinewood derby. At the same time, Barry has reasons to prefer his kid to win the pinewood derby, his neighborhood to win a soccer tournament and his company to secure a contract. This is the standard situation of an inter-personal moral dilemma, which is accounted for by C but not by M. In more general terms, under normal circumstances there is no room for morality to pick a winner among two competitors in chess—this is because chess competitions should normally be decided by the criterion of who is the best in playing chess, not on the moral grounds. Yet, there is an agent-relative objective moral reason for family and close friends of each competitor to root for those with whom they have special moral bounds (and it is *sometimes* morally wrong for them to remain impartial, or to root for the opposite team). But in this paper we just need to say that a moral theory that adopts such moral reasons is free of the charge of inconsistency.

Moral theory C, which is able to account for the conflict of interests to which there is no morally prescribed solution, is like an umpire who partakes in a game only in so far as to enforce the rules such as '*Jack's family should help Jack win and Mike's family should help Mike, and they should respect certain constraints in doing so.*' The balance of reasons, including moral reasons, can be viewed as different from the point of view of one's company, friend or a family member, although the good of other people affected by their actions (or failures to act) should also remain on the moral picture.<sup>10</sup> Hence, it is morally wrong to maximize only goodness *sub specie aeternitatis*.

Moral agents not only ought to follow an agent-neutral constraint that fair play ought to be followed by everybody, but also an agent-

---

<sup>10</sup> This is how moderate patriotism differs from blind chauvinism which does not take into account the interests of non-members of the fatherland community at all.

relative one that every person must root for their respective team, at least in certain instances. The umpire's role is *emphatically not to pre-select the winners and losers*, and the role of moral theory in normal competitive situations is not always to pick the so-called 'morally preferred outcomes'; this feature of consequence-based moral evaluation is more similar to deontology than to standard consequentialism. C-ethics is comfortable with competition, including zero-sum games, which are an important part of life. The structure of moral theory proposed by Sen accommodates agent-relativity required by this model.

#### 4. Conclusions

We must agree with Parfit that C does not provide consistent outcome-related moral recommendations since it allows situations in which the best moral interests of various agents conflict. Yet we must also agree with Sen that any moral theory that provides such recommendations is descriptively deficient, since it may be hard to apply or simply lack fit with the reality. Human beings do face inter-personal zero-sum games in life. A moral theory that tries to pre-judge the outcomes of such competitive situations by adopting an archangel's viewpoint refuses to note this fact, or rather views it as inadmissible to the normative moral theory. A better kind of moral theory acknowledges that there are different agent-relative rankings from the viewpoint of various agents and that such rankings should guide moral evaluations.

Sen's agent-relative broad consequentialism provides a framework for such moral theory. Hence, we demonstrate that such theory is possible; we also point to some of the reasons to believe that it is plausible. Since agent-relativity of moral reasons leads to inter-personal moral conflicts, the role of the moral system does not seem to be to always solve or dissipate such conflicts up. Sometimes, all that moral theory can do is to provide the rules of fair competition among various reasons, interests and moral concerns.

#### References

- Alexander, J.M. (2003). Evolutionary Game Theory. In *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/game-evolutionary>, Feb. 20, 2007.
- Archard, D. (1995). Moral Partiality. *Midwest Studies in Philosophy*, 20, 129–41.
- Benhabib, S. (1986). The Generalized and the Concrete Other: The Kohlberg-Gilligan Controversy and Feminist Theory. *Praxis International*, 5/4, 402–24.
- Blum, L. (1988). Gilligan and Kohlberg: Implications for Moral Theory. *Ethics*, 98, 472–91.

- Boltuć, P. (2001). Moral Neighborhoods. *Dialogue and Universalism*, 11, 117–33.
- Boltuć, P. (2005). Moralność a stronniczość. *Analiza i Egzystencja*, 2, 25–38.
- Boltuć, P. (2008). Inherent Second-Person Moral Reasons. In J. Abbarno (Ed.), *Inherent Instrumental Reasons*, in publication.
- Cottingham, J. (1986). Partiality, Favouritism and Morality. *The Philosophical Quarterly*, 36, 357–73.
- Cottingham, J. (1988). The Ethical Credentials of Partiality. *Proceedings of the Aristotelian Society*, 1–21.
- Dancy, J. (1983). Ethical Particularism and Morally Relevant Properties. *Mind*, 92, 530–47.
- Dancy, J. (1993). *Moral Reasons*. Oxford: Blackwell.
- Dancy, J. (2004). *Ethics without Principles*. Oxford: Oxford University Press.
- Duska, R., & Whelen, M. (1975). *Moral Development: A Guide to Piaget and Kohlberg*. New York: Paulist.
- Elster, I. (1979). *Ulysses and the Sirens*. Cambridge: Cambridge University Press.
- Elster, I. (2002). *Ulysses Unbound: Studies in Rationality, Precommitment, and Constraints*. Cambridge: Cambridge University Press.
- Flanagan, O., & Jackson, K., (1987). Justice, Care, and Gender: The Kohlberg-Gilligan Debate Revisited. *Ethics*, 97, 622–37.
- Friedman, M. (1991). The Practice of Partiality. *Ethics*, 101, 818–35.
- Friedman, M. (1993). *What Are Friends For? Feminist Perspectives on Personal Relationships and Moral Theory*. Ithaca, NY: Cornell University Press.
- Gilligan, C. (1982). *In a Different Voice*. Harvard University Press.
- Hampton, J. (1998). *The Authority of Reason*. Cambridge: Cambridge University Press.
- Hare, R. M. (1981). *Moral Thinking*. Oxford: Clarendon Press.
- Harman, G. (2005). Moral Particularism and Transduction. *Philosophical Issues*, 15, 44–55.
- Kirk, R. (1985). *The Conservative Mind*. Washington: Regnery Publishing.
- Kohlberg, L. (1981). *Essays on Moral Development*, vol. 1, *The Philosophy of Moral Development: Moral Stages and the Idea of Justice*. San Francisco: Harper and Row.
- Kohlberg, L. (1982). A Reply to Owen Flanagan and Some Comments on the Puka-Goodpaster Exchange. *Ethics*, 92, 513–28.
- Kohlberg, L. (1984). *Essays on Moral Development*, vol. 2, *The Psychology of Moral Development*. San Francisco: Harper and Row.
- McClennen, E. (1990). *Rationality and Dynamic Choice*. Cambridge: Cambridge University Press.
- McNaughton, D. (1988). *Moral Vision: An Introduction to Ethics*. Oxford: Blackwell.
- Nagel, Th. (1986). *The View from Nowhere*. Oxford: Oxford University Press.
- Nagel, Th. (1991). *Equality and Partiality*. Oxford: Oxford University Press.
- Parfit, D. (1979). Is Common-Sense Morality Self-Defeating. *The Journal of Philosophy*, 76, 533–45.
- Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- Pargetter, R. (1991). Kinship and Moral Relativity. *Philosophia*, 20, 345–61.
- Porter, M. (1998). Porter on Competition. *Harvard Business Review Book*. Harvard Business School Press.
- Portmore, D.W. (2003). Position-Relative Consequentialism, Agent-Centered Options, and Supererogation. *Ethics*, 113, 303–32.
- Portmore, D. (forthcoming). Consequentializing Moral Theories. *Pacific Philosophical Quarterly*.

- Regan, D. (1983). Against Evaluator Relativity: A Response to Sen. *Philosophy and Public Affairs*, 12, 93–112.
- Scanlon, Th. (2000). *What We Owe to Each Other?*. Cambridge: The Belknap Press of Harvard University Press.
- Scheffler, S. (ed.) (1988). *Consequentialism and Its Critics*. Oxford: Oxford University Press.
- Scheffler, S. (1997). Relationships and Responsibilities. *Philosophy and Public Affairs*, 26, 189–209.
- Scheffler, S. (2001). *Boundaries and Allegiances*. Oxford: Oxford University Press.
- Sidgwick, H. (1874). *The Methods of Ethics*. London: Macmillan.
- Sen, A. (1982). Rights and Agency. *Philosophy and Public Affairs*, 11, 3–39.
- Sen, A. (1983). Evaluator Relativity and Consequential Evaluation. *Philosophy and Public Affairs*, 12, 113–32.
- Sen, A. (1993). Positional Objectivity. *Philosophy and Public Affairs*, 22, 126–45.
- Sen, A. (2000). Consequential Evaluation and Practical Reason. *The Journal of Philosophy*, 97, 477–502.
- Tatarkiewicz, W. (1976). *Analysis of Happiness*. Hague: Martinus Nijhoff.
- Tatarkiewicz, W. (2007). Happiness and Time. In S.M. Cahn & Ch. Vitrano (Eds.), *Happiness Classic and Contemporary Readings in Philosophy*. Oxford: Oxford University Press.
- Thomas, A. (2007). Values, Reasons and Perspectives. University of Kent, <http://www.logical-operator.com/VRP.pdf> , Feb. 20, 2007.