Just consequentialism and computing

James H. Moor
Department of Philosophy, Dartmouth College, NH, USA. E-mail: James.H.Moor@Dartmouth.edu

Abstract. Computer and information ethics, as well as other fields of applied ethics, need ethical theories which coherently unify deontological and consequentialist aspects of ethical analysis. The proposed theory of just consequentialism emphasizes consequences of policies within the constraints of justice. This makes just consequentialism a practical and theoretically sound approach to ethical problems of computer and information ethics.

Introduction

The malleability of computers allows them to be used in novel and unexpected ways, ways for which we frequently do not have formulated policies for controlling their use.[1] Advancing computer technology produces policy vacuums in its wake. And even when relevant policies do exist, they are often ambiguous or inadequate as they were designed for times with a less versatile technology than computing. A basic job of computer ethics is to identify these policy needs, clarify related conceptual confusions, formulate appropriate new policies, and ethically justify them.

Policies are rules of conduct ranging from formal laws to informal, implicit guidelines for action. Policies recommend kinds of actions that are sometimes contingent upon different situations. "Turn off your computer when you are finished" is a policy though probably one without much ethical significance. "Don't steal computer chips" is another policy with more obvious ethical content. Even when policies are the policies of others they can help us to regulate our lives. We know what to expect and can react accordingly. If a restaurant has the policy of using caller ID to capture the numbers of incoming phone calls, then, if we don't want to have our phone number known by the restaurant, we can block the caller ID system or use another restaurant. In this discussion our concern is with those computing policies that make an ethical difference and how to evaluate them properly.

Viewing issues in the ethics of computing in terms of policies is important. Policies have the right level of generality to consider in assessing the morality of conduct. Every action can be considered as an instance of a policy – in this kind of situation such and such action is allowed or required or forbidden. Understanding actions as exemplars of more general normative prescriptions promotes responsibility and more careful reflection on conduct. The word "policy" is intended to suggest both that there can be justified exemptions (policies are not absolute rules) and a level of obligation (policies are not mere suggestions).

We want our policies for computing to be ethical, but what should we look for when constructing ethical computing policies? When we turn to traditional ethical theories for help we discover a strong rivalry exists between the leading contenders – consequentialist theories that emphasize the consequences of actions and deontological theories that stress rights and duties. Indeed, consequentialist theories and deontological theories are often presented as hopelessly incompatible. Philosophers, who perhaps take some pernicious delight in being gadflies and stirring things up, sometimes to revel in the conflicts among ethical theories. But the absence of resolution among the ethical theories leaves many with a somewhat jaundiced estimate of the value of ethical theory altogether. Applied ethicists, searching for practical guidance, find themselves immersed in ad hoc analyses of ethical problems and selecting solutions from an inconsistent pile of principles.

I believe that ethics needs more unifying theories that call upon the various strengths of the traditional approaches to ethics. One is reminded of the story of the elephant in which each observer obtaining evidence from only part of the elephant gives a description of the creature that diverges dramatically from the descriptions by others. Of course, there is an overall description of the elephant that makes sense of the individual apparently incompatible descriptions. Similarly, the elephant of ethical theory has been described by conflicting descriptions. Our job is to

---

[1] James H. Moor. What Is Computer Ethics? *Metaphilosophy*, 16(4), pp. 266–275, 1985.

try to discover, if possible, an overall description of the elephant that will make sense of these apparently incompatible descriptions. This paper takes a few steps in that direction.

Consequentialism constrained by justice

The ethical evaluation of a given policy requires the evaluation of the consequences of that policy and often the consequences of the policy compared with the consequences of other possible policies. If our actions involving computers had no harmful consequences, policies would not be needed. However, conventional consequentialism has well-known shortcomings. Among other objections consequentialism seems to be insensitive to issues of justice. I believe there may be a unifying ethical theory that allows us to take into account the consequences of policies while at the same time making sure that these policies are constrained by principles of justice.

When considering consequences we evaluate the benefits and harms. Human beings have a common nature. At the core humans have similar kinds of values, i.e., what kinds of things they consider to be goods (benefits) and what kinds of things they consider to be evils (harms). In general the core goods include life, happiness, and autonomy and the core evils include death, unhappiness, and lack of autonomy. By 'happiness' I mean simply 'pleasure and the absence of pain'. The notion of autonomy here needs some explanation as the term is used by others in different ways. Obviously, humans do not share all their goals in common. But no matter what goals humans seek they need ability, security, knowledge, freedom, opportunity and resources in order to accomplish their projects. These are the kinds of goods that permit each us to do whatever we want to. For brevity I will call this set of goods the 'goods of autonomy' or simply 'autonomy'. The goods of autonomy are just the goods we would ask for in order to complete our projects.[2] For a given individual the goods of autonomy are not necessarily any less valuable than happiness or even life. Some people will give up happiness for knowledge and some will give up life without freedom. Individuals will rank the core values differently and even the same individual may rank the core values differently during her lifetime. But for good evolutionary reasons all rational human beings put high positive value on life, happiness, and autonomy at least for themselves and those for whom

they care. If they did not, they would not survive very long.

Of course, humans are not necessarily concerned about the lives, happiness, and autonomy of others, but they are concerned about their own. To be ethical one must not inflict unjustified harm (death, suffering, or decreased autonomy) on others. To take the ethical point of view is to be concerned for others at least to the extent that one tries to avoid harming them. The fact that humans value and disvalue the same kinds of things suggests that, contrary to the claims of some kinds of relativism, there may be common standards by which humans of different cultures can evaluate actions and policies.[3]

The combined notions of human life, happiness, and autonomy may not be far from what Aristotle meant by "human flourishing". Thus, from an ethical point of view we seek computing policies that at least protect, if not promote, human flourishing. Another way to make this point is to regard the core goods as marking fundamental human rights – at least as negative human rights. Humans ought to have their lives, happiness, and autonomy protected. And this principle of justice – the protection of fundamental human rights – should guide us in shaping ethical policies for using computer technology. When humans are using computer technology to harm other humans, there is a burden of justification on those doing the harming.

We have the beginning of a unifying ethical theory. When evaluating policies for computing, we need to evaluate the consequences of the proposed policies. We know what the core goods and evils are. And we want to protect human rights. Nobody should be harmed. The theory so far does constrain consequentialism by considerations of justice but perhaps too well. Realistically, harmful consequences cannot always be avoided and sometimes it seems justified to cause harm to others, e.g., in giving punishment or defending oneself. Is there an approach to justice that will allow us to resolve conflicts of action or policy when causing harm seems unavoidable or even reasonable?

Bernard Gert provides us with a notion of moral impartiality that offers a good insight to justice that is useful in resolving these conflicts.[4] His moral theory inspires and informs the following analysis though I will not do justice to his account of justice.

Justice requires an impartially toward the kinds of policies we allow. Therefore, it is unjust is for

---

[2] 'ASK FOR' is a good way to remember the goods of autonomy: [A]bility, [S]ecurity. [K]nowledge, [F]reedom, [O]pportunity, [R]esources.

[3] James H. Moor. Reason, Relativity, and Responsibility in Computer Ethics. *Computers & Society*, 28(1), pp. 14–21, 1998.

[4] Gert, Bernard. *Morality: Its Nature and Justification.* Oxford, Oxford University Press, 1998.

someone to use a kind of policy that he would not allow others to use. Consider the policy of a company knowingly and secretly installing defective computer chips in its products for sale. This policy of installing defective chips in products is an instance of a more general kind of policy that permits the manufacture and deceptive selling of resources that may significantly harm people. No rational person would accept this kind of public policy, for that would have to allow others to follow the same policy putting oneself at an unacceptable level of risk.

Now consider another example of a policy. A person who intrudes into my house uninvited will be electronically monitored and reported electronically to the police. Such a policy will result in causing harm to an intruder. Is it justified? Violating my security is a harm to me; it violates my privacy and puts me at risk. If we take an impartial point of view, then the policy stated more abstractly becomes people are allowed to harm (within limits) other people who are unjustly harming them. A rational, impartial person could accept such a policy in that others could be allowed to follow it as well. Others following such a policy will not harm you unless you harm them first.

This impartiality test can be used to evaluate whether or not exceptions to existing policies are justified. Suppose an airline has a policy of flying its planes on time and allowing computer systems to do the flying. This permits efficient and timely service on which its customers count. Now suppose the airline suddenly discovers a bug in their software that may endanger its planes and its passengers. Clearly the airline should request that human pilots fly its planes already in the air and ground the rest until the software problem is located and removed. Passengers will be harmed by delays and lost opportunities to fly but any rational, impartial person would allow such an exception to this policy if it meant avoiding significant risk of death. Such an exception can be made part of the policy and publicly advocated.

Gert applies his account of impartiality to potential violations of moral rules. This is a two step procedure in which one abstracts the essential features of the situation using morally relevant features and then asks whether the resulting rule so modified could be publicly allowed, i.e., what consequences would follow if everyone knew about this kind of violation and could do it. I have done an analogous operation on the above examples by abstracting the policy and then asking what would be the consequences if such polices were publicly allowed.

Gert refers to his view of impartiality as 'the blindfold of justice'. The blindfold removes all knowledge of who will benefit or will be harmed by one's choices.

This is similar to John Rawls' veil of ignorance[5] except that Gert allows those who are blindfolded to assign different weights to the benefits and harms. As a result there is room for disagreement that is not permitted in Rawls' account. Not all rational, impartial people will agree on every judgment. However, some judgments will be shared by all rational, impartial people.

If the blindfold of justice is applied to computing policies, some policies will be regarded as unjust by all rational, impartial people, some policies will be regarded as just by all rational, impartial people, and some will be in dispute. This approach is good enough to provide just constraints on consequentialism. We first require that all computing policies pass the impartiality test. Clearly, our computing policies should not be among those that every rational, impartial person would regard as unjust. Then we can further select policies by looking at their beneficial consequences. We are not ethically required to select policies with the best possible outcomes, but we can assess the merits of various policies using consequentialist considerations and we may select very good ones from those that are just.

The good as the enemy of the just

We should develop computing policies in such a way that they are above all just. Then we can make the policies as good as reasonably possible. Our first priority should be to avoid unjustifiable harming of others by protecting their rights and then to increase benefits.

It may be tempting in some situations to focus on the strikingly good consequences of a policy while ignoring injustice. The potential good in given situation is so good that it seems to justify being unjust. Suppose a company wants to market a set of CDs with extensive personal information about people in a country.[6] The good from a marketing point of view is staggering. Every marketer would be able to know considerable personal details about the citizens of a country and send out relevant marketing materials to those who need it. The good that looks so good in the short run may be overwhelmed by harm in the long run. The conventional consequentialist would point out that harm is done initially to people's autonomy and should not be overlooked. Because people's privacy is invaded, their security is reduced. With release of the

[5] John Rawls. *A Theory of Justice*, Cambridge, Massachusetts, Harvard University Press, 1971.

[6] In 1990 Lotus Corporation was planning to sell software with a database of 120 million Americans including names, addresses, incomes, buying habits, and other personal data. Lotus aborted the idea due to public outrage.

CDs people would become vulnerable to misuses of the information such as losing employment or losing insurance privileges. The just consequentialist has a further concern. Even if we stipulate that no such long term harm will occur through the release of this information, there is still collateral damage. If releasing these CDs filled with personal information is allowable, then similar activities are allowable in similar situations. In other words, by our impartiality principle anyone else can inflict similar harms and put people at similar risks if the good from doing so seems substantial. Given the fallibility of humans, no rational, impartial person would be willing to take this risk.

Consider another example. The current copyright law protects software. Suppose someone decides to give a copy of her word processing software illegally to another person, a lowly graduate student.[7] There is potentially lost revenue to the software manufacturer but suppose that the graduate student receiving the software is living on limited means and would not be able to buy the software anyway. The student is someone who would benefit greatly from having the software if she had it. Why not illegally copy software in such cases? The foreseeable good is tempting. The conventional consequentialist may agree or more cautiously respond that the graduate student or the dispenser of illegal software may be discovered which would lead to unanticipated harm. For example, the university protecting itself from lawsuits may have a rule that a doctoral dissertation discovered to be written on illegal software is not acceptable. The just consequentialist would point out in addition that there is collateral damage. If someone can violate the law in this case to benefit a friend, then by the impartiality principle others can do so as well. Not only may other software be copied illegally to help friends, but other laws seem equally open to violation on the presumption that some people are being helped. Given the fallibility of humans, no rational, impartial person would want to take this risk.

This analysis is based upon the assumption that the copyright law itself is just. The law has been properly enacted and does not unjustifiably violate anyone's fundamental human rights. The copyright law does seem to be just in this way. However, this leaves open the question whether the copyright law could be better. We might in fact want to allow greater fair use in the copying of software and enact better laws.

Sometimes it is said, "The ends do not justify the means." In one sense this statement is clearly false. If the ends don't justify the means, what would? If we take the ends to be our core goods, then they are

---

7 Compare Nissenbaum (1995) and Gert (1999).

satisfactory ends for the purposes of justification. In another sense this claim may mean "The ends do not justify any means that harm people." This is also false. One has to look at a situation impartially and ask what kinds of policies for conduct should apply. Sometimes harming some people somewhat to avoid much greater harm to them or others is completely justified. Or the claim might mean "The ends do not justify using unjust means." This is the interpretation of the claim that is true. This is precisely what happens when the good becomes the enemy of the just. Good ends somehow blind us to the injustice of the means.

We want good computing policies that promote human flourishing, consequences are important, but only as long as the policies themselves remain just. Unjust policies will in the long run, both directly and indirectly, undermine the benefits of these policies no matter how good they are.

Computing in uncharted waters

Setting ethical policies for computing might be compared to setting a course while sailing. A sailor may chart a course to her destination by dead reckoning carefully laying out the course in a straight line on a chart. But sometimes there are no charts, and, even if there are, experienced sailors know how difficult it is to keep the course true. Winds, currents and tides are constantly changing. Sailors do not continually steer on course and trim their sails precisely. Midcourse adjustments are necessary and proper and should be expected. Similarly, setting ethical policies for computing is a something of an approximation. Nobody can accurately predict the many changing factors in computing situations. Given the logical malleability of computing many new opportunities and unexpected developments will arise. Human reactions to these new possibilities are equally difficult to predict. Midcourse adjustments in computing policy are necessary and proper and should be expected.

Sailors take danger bearings – to avoid dangerous objects such a reef. Certain courses should not be taken. This leaves open many other courses as good options, and reasonable sailors may disagree about which is best. Some may prefer to get to the destination in the shortest time, others may want to see a scenic island, and still others may wish to set a course to improve a suntan. Similarly, in setting computing policy there are policies we want to avoid. We do not want our policies to violate fundamental human rights unjustly. But given that our policies are just, many good policy options may be available though people may have legitimate disagreements about which is best.

References

Bernard Gert. *Morality: Its Nature and Justification*. Oxford, Oxford University Press, 1998.

Bernard Gert. Common Morality and Computing. *Ethics and Information Technology* 1 (1): 57–64, 1999.

James H. Moor. What Is Computer Ethics? *Metaphilosophy* 16 (4): 266–275, 1985.

James H. Moor. Reason, Relativity, and Responsibility in Computer Ethics. *Computers & Society* 28 (1): 14–21, 1998.

Helen Nissenbaum. Should I Copy My Neighbor's Software? Deborah G. Johnson and Helen Nissenbaum, editors, *Computers, Ethics, and Social Values*, Englewood, New Jersey, Prentice-Hall, pp. 201–213, 1995.

John Rawls. *A Theory of Justice*. Cambridge, Massachusetts, Harvard University Press, 1971.