THE CAMBRIDGE HANDBOOK OF

INFORMATION AND COMPUTER ETHICS

Edited by Luciano Floridi

4 The use of normative theories in computer ethics

Jeroen van den Hoven

4.1 Introduction

Without Information and Communication Technologies (ICTs) many of the activities that we undertake in the twenty-first century in the world of trade, finance, transport, healthcare, science, education, administration, management, communication, energy supply, industrial production, defence, engineering and technology would be impossible. Computers have become a necessary condition for all of our large-scale projects and complex endeavours.

Some of the major moral problems of Information Societies at the beginning of the twenty-first century concern the quality and reliability of information, control and governance of the Internet, responsibility for data processing, property of software and privacy and protection of personal data and the quality of life. There are also problems concerning power and dominance of commercial parties, equal access and fair distribution of information. A relatively new set of issues concerns the way the technology invades our daily lives and affects the moral development of children and young people who have had long and intense exposure to the technology and the content it offers. This listing is not exhaustive and new issues are constantly appearing as the technology develops. The issues occupy a prominent place in public debates, demand attention in the policy arena and usually require regulation because the lives and interest of many are potentially affected. Computer and information ethics has tried to shed light upon these and other issues in the last decades.[1]

ICTs have properties which make it difficult to make up our minds concerning the answers to the moral questions to which they give rise and it is certainly not the type of technology that we can decide to turn off or jettison should we become uncomfortable with its problems and results. ICTs are (1) *ubiquitous* and *pervasive* in a way in which our most common technical artefacts are not. Common household appliances and ordinary objects nowadays *are* computers and will often be interconnected through wireless network

---

[1] See for overviews Himma and Tavani 2008, Johnson 2009, van den Hoven and Weckert 2008, Weckert 2007.

technologies. More and more everyday objects and artefacts are woven into an Internet of Things that eventually meshes with the Internet of People. More and more tasks involve interaction with computers or computerized tools and devices. Technology and infrastructure which is omnipresent has a tendency to blend into the background, become translucent and disappear from our radar screen, making it more difficult to assess its role (Bowker and Star 1999). (2) ICTs are a *universal technology*, because of their 'logical malleability' (Moor 1985). Digital computers are in essence Turing Machines that can be used to simulate, communicate, recreate, calculate, and so much more, in all domains of life in all sectors of society. The entities manipulated on the machine level can be made to stand for everything that can be articulated and expressed in terms of symbols. We can use the same machine to simulate a weather storm, to distribute electrical power in a part of the country, to run a production plant and archive government information. It is therefore often difficult to see the common elements in the many manifestations and applications of ICTs. ICTs are (3) a *meta-technology*, that is, a technology which forms an essential ingredient in the development and use of other technologies. It helps us to drive cars, make medical images, produce petrol and distribute goods over the world. This may obscure the fact that problems which are identified with the first-order technology are, in fact, problems with the meta-technology. ICTs are also (4) a *constitutive technology*. Computing technology co-constitutes the things to which it is applied. ICTs are often characterized as an *enabling* technology and it is certainly correct to say that they enable us to do new things and to do old things in new ways, but this must hide the fact that, where ICTs are introduced they *transform* our old practices, discourses, relations and our experiences in fundamental ways and they are partly constitutive of new practices. If they are used in health care, health care will change in important ways, if they are used in science and education, science and education will never be the same again, if Internet and the World Wide Web are introduced in the lives of children, their lives will be very different from the childhood of people who grew up without online computer games and social networking sites. Furthermore, ICTs are about *information*[2] (5). Information is so important to human beings that we tend to forget that we use and process information constantly; we need it in deliberation, planning, choice, decision-making, preference formation and judgement. If information is inaccessible, wrong, inaccurate or incomplete, the results of these cognitive processes are compromised. ICTs provide the mechanisms to channel and manipulate this all-important good, hence the moral significance of their

[2] Luciano Floridi's work forms a broad-ranging and in-depth study of this aspect, see his contribution 'Information Ethics: Its Nature and Scope' in van den Hoven and Weckert 2008, pp. 40–66, and the special issue of *Ethics and Information Technology*, vol. 10, 2008, nos 2–3.

evaluation, regulation and design. ICTs are also the expression of prior choices, norms, values, and decisions (6). ICT applications are not neutral, but contain the values and norms of those who have designed and engineered them. An abundance of research provides evidence of intentional or inadvertent incorporation of norms in software (Friedman 1997). Finally, ICTs revolve around new entities, such as digital computers, software and information goods, which give rise to new practices and experiences. This makes it sometimes difficult to account for them in terms of traditional moral and legal views (7).

These characteristics taken together form an explication of the common observation that ICTs play a central but confusing role in our lives. It is often not immediately clear that ICTs merit special attention and require moral evaluation and analysis of the sort that computer and information ethics attempt to provide.

A safe starting point for moral thinking is to look simply at the effects the new entities have on people, the environment and on everything we endow with moral standing, what people can do to each other by means of these entities, how they constrain or enable us, how they change our experiences and shape our thinking, how they affect our relationships and balances of power. Another starting point is to turn to some of the ethical theories in the history of philosophy, such as utilitarianism, Kantian ethics or virtue ethics, and see whether they can shed light on the problems. This is what computer ethics has done in the past three decades. This is also how we proceeded in the case of thinking about the car, the television and the atom bomb when they were introduced, and this is how we shall proceed in the case of evaluating brain imaging technology and the use of carbon nano-tubes, artificial agents and the applications of advanced robotics. We certainly need to retain what is obviously helpful in traditional ethical thinking as it applies to ICTs, but a fully adequate ethical treatment of ICTs in the decades ahead requires a somewhat different approach to moral theorizing from the ones that have been tried thus far.

First of all, there is no other way for moral thinking in the field of ICTs than to embrace a robust *conceptual and value pluralism* – which does not imply moral scepticism or moral relativism (4.2). Secondly, the conception of ethical theory or ethical thinking must accommodate the pluralist condition and be *empirically informed, realistic and practical,* so as to provide guidance and direction in cases where information technology is actually used (4.3). Thirdly, it should support *conceptual reconstructions* of ethical key concepts that play an important role in the discourse that is actually used in the description, evaluation and shaping of the technology, in order to fill conceptual vacuums as described by Moor (4.4). Finally, it should focus on issues of moral *design* of ICT applications at an early stage of development and not only focus on their *evaluation ex post* (4.5).

4.2 Value pluralism

Christine Korsgaard has pointed out that 'one of the most important attributes of humanity is our nearly bottomless capacity for conferring value on most anything. It is not because of our shared values that we should accord consideration to one another but because of our shared capacity for conferring value. In other words, that fact about human nature is part of what makes liberal democratic forms of the state the right ones' (Korsgaard 2003, p. 73). This fact about human beings and human lives has served as a point of departure of much of contemporary moral theorizing. We confer value on different things, but we also confer different values on one and the same thing.

Since Isaiah Berlin wrote his *Two Concepts of Liberty* (Berlin 1958), many leading contemporary philosophers working in a broadly liberal tradition have subscribed to the idea that there are many different and incommensurable – or at least *de facto* conflicting – values or sources of moral evaluation (Galston 2002, pp. 3–15). In a different context, Berlin used an ancient proverb about the difference between the fox and the hedgehog to illustrate the difference between monists and pluralists: the fox sees many small things, the hedgehog sees one big thing (Berlin 1957). Many contemporary moral philosophers see many small things instead of one big thing when looking closely at ethics and morality: Bernard Williams, Thomas Nagel, Martha Nussbaum and Amartya Sen, John Rawls, Joseph Raz, Robert Audi (2007), James Griffin (1996) – to name a few of the most prominent – all defended forms of value pluralism.

The pluralist position is paradigmatically exemplified in Thomas Nagel's seminal paper 'The Fragmentation of Value'.[3] Nagel states there that he does not believe that 'the source of value is unitary...I believe that value has fundamentally different kinds of sources and that they are reflected in the classifications of values into types.'[4] Nagel distinguishes five fundamental types of value: Utility, General Rights, Special Obligations, Commitments to Own Projects and Perfectionist Ends.[5] Human lives, endeavours and social relationships are variegated and intricate. The problems with which persons are confronted are multifarious, their actions have multiple ramifications and a range of effects upon others. People can see things from radically different perspectives. They can look at results of their actions and at the springs of their actions, they can look at things from their particular point of view and they can identify and sympathize with others close to them, or with distant others. They can look at their own situation with a 'view from nowhere' (Nagel 1986), or they can look at the Universe from their personal point of view and they can switch between these perspectives, without feeling that one perspective is more real or more important than the other. These points of view and valuing

---

[3] Thomas Nagel 'The Fragmentation of Value', reprinted in Gowans 1987, pp. 174–187.
[4] Thomas Nagel 'The Fragmentation of Value', reprinted in Gowans 1987, p. 177.
[5] Thomas Nagel 'The Fragmentation of Value', reprinted in Gowans 1987, p. 175.

are equally valid *all other things being equal.* It therefore can not be the case that the only thing which counts from the moral point of view is consequences or outcomes and the maximization of utility, happiness, pleasure or money. Nor can it be the case that compliance with one formal moral principle of duty and human dignity can be the only right making criterion, whatever the consequences. Or that the special obligations and loyalties that one has because of one social role or position in a social network are always all-important and trump all considerations of utility or general rights. A person's commitment to his or her own personal projects certainly also counts for something in cases of conflicts with the maximization of overall utility, general rights or special obligations to significant others, but for how much must be determined in every case anew. Even an appeal to perfectionist values regarding how an ideal or perfect human being ought to behave, e.g. regarding sexual matters or personal hygiene and aesthetics, may have some initial plausibility, but are certainly overruled in cases of conflicts with general rights or utility.

Different normative ethical theories and traditions have singled out one type of value, epitomized it and have consequently downplayed the importance of the others, reduced them to their value of choice, or have eliminated them altogether. Monistic views of moral theory presuppose that all one needs to know is one value or one simple principle which expresses it. To believe that there is one master value that trumps all others – whether it is human dignity or the maximization of utility, self-interest or human flourishing – amounts to an unduly narrow view of the complexity of moral problems and the human condition, which ought to be avoided, especially in applied ethics, which aspires to be relevant to technology assessment and public policy making.

Another dimension of the robust pluralism referred to above is *conceptual* pluralism in ethical theory. Wittgenstein remarked that 'mathematics is a motley', which led Hilary Putnam to characterize ethics as 'a motley squared' and to observe that 'philosophers who write about the subject so often ignore vast tracks of ethical judgment' (Putnam 2004, p. 72). Ethics may be about praise and blame, about evaluation or prescription, action guidance, conflict resolution, about virtues and character traits, about the logic of obligation and permission, about human rights, about basic needs, utility, outcomes and consequences, money, well-being, norms, principles, ideals, capabilities, responsibilities, duties, interest and preferences and values. It may be about highly general or universal truths or about context-specific considerations. Depending on the situation, we may want to utilize any of these concepts and vocabularies. To foreclose the use of them with their associated background views in favour of one seems unduly restrictive and reductive in practical matters.

In discussions on privacy online, for example, we may sometimes want to express the importance of privacy in terms of individual autonomy or freedom, then in terms of intimacy and personal relationships, basic needs,

human rights, in terms of fiduciary duties of professionals, and responsibilities of management, the logical structure of a policy document, the subjective expected utility – costs and benefits – of a proposed set of regulations. Our societies are complex, information technology is complex and hence the privacy issue is complex. Under the heading of privacy violations, a variety of moral wrongdoings belong, such as physical assault, theft, discrimination, economic disadvantage and loss of moral autonomy (van den Hoven in van den Hoven and Weckert 2008). We need access to the relevant vocabularies and background views to articulate and assess the range of wrongs and think about the best ways to prevent them.

This amounts to what Hilary Putnam has called *Pragmatic or Conceptual Pluralism*, which recognizes that 'in everyday language we employ many different kinds of discourses, discourses subject to different standards and possessing different sorts of applications, with different logical and grammatical features', and which denies that there could be one sort of language game sufficient for the description of all of (moral) reality (Putnam 2004, pp. 21, 48 ff.).

Jim Moor's 'Core Value Approach' to computer ethics is a paradigmatic example of Value Pluralism applied to Computer Ethics.[6] Moor identifies moral values such as *life, health, happiness, security, resources, opportunities* and *knowledge* which are vital to the survival of any community, and claims that all communities do in fact value them. Indeed, if a community did *not* value the 'core values', it soon would cease to exist. Moor used 'core values' to examine computer ethics topics like privacy and security and to add an account of justice, which he called 'just consequentialism' which combines 'core values' and consequentialism with Bernard Gert's deontological notion of 'moral impartiality'.[7]

4.3   Moral theory

4.3.1   Primacy of practice

Ethics is a department of practical philosophy and thus primarily concerned with practical problems and action. The aim of moral argumentation, moral reasoning and judgement is the rational justification and settlement of disagreement and conflicts about who one wants to be, what to do, the constraining of self-interest and the fostering of cooperation and peaceful coexistence of sentient creatures in a shared habitat. Moral thinking points to reasons for constraining self-interested behaviour and self-serving strategies. We reflect upon and attempt to improve our moral beliefs and ideas with the end in view of finding answers to the question how to lead a flourishing life, how to act,

---

[6] Moor (2001) in Spinello and Tavani 2001, pp. 98–105.
[7] Moor (2001) in Spinello and Tavani 2001, pp. 98–105.

decide and choose in such a way as to pursue our own happiness without interfering with the similar pursuit of others.

The main aim in ethics is not to establish a general theory and a set of eternal truth but to provide reasoned solutions and clarifications to practical problems. We engage in moral theory in order to create the intellectual resources that can help us to determine which of our moral beliefs are worthiest of our endorsement. Dewey thought that 'Philosophy recovers itself when it ceases to be a device for dealing with the problems of philosophers and becomes a method, cultivated by philosophers, for dealing with the problems of men.'[8]

4.3.2   Theoretical pluralism

Value pluralism has implications for an account of ethical theory. First, different theories and their associated core values may capture different morally relevant aspects of a particular case, without necessarily leading to unique and correct answers. Secondly, since there are different vocabularies and conceptual frameworks available for describing situations, each of them may give a different answer to questions of salience and relevance and may even lead to the articulation of different sets of moral questions.

The individuation and description of concrete cases has implications for what is subsumed under a general moral rule, principle or theory. This is known as the problem of relevant description. Anscombe observed that an act-token will fall under many possible principles of action, which makes it difficult to tell which act description is relevant for moral assessment (Anscombe 1958). Should we, Onora O'Neill asks in her discussion of Anscombe's problem, 'assess an action under the description that an agent intends it, or under descriptions others think salient, or under descriptions that nobody has noted' (O'Neill 2004, p. 306)? And how do we evaluate the actions of persons who – according to us – fail to see the morally significant descriptions of what s(he) does? Bernard Gert gives an example of how the description of the case is also of crucial importance in computer ethics (Gert 1999). He analyses Nissenbaum's analysis of moral permissibility of copying software for a friend. Gert remarks that disagreement about this issue may be due to the fact that one of the partners to the disagreement has too narrow a description of the kind of violation to launch ethical thinking in the right direction. Some may describe it as 'helping a friend', some as 'illegally copying a software program', or as 'violating a morally acceptable law to gain some benefit'. On the basis of the latter description, Gert claims that 'no impartial rational person would publicly allow the act' (p. 62). As we can see from this example, pluralism does not imply that it is impossible to argue on good grounds that particular

8 Dewey, quoted in Putnam 2004, p. 31.

proposals and arguments are better than others. Pluralism does not imply moral scepticism or moral relativism.

Anti-theorists, to which Stuart Hampshire, John McDowell, Annette Baier, Bernard Williams and Martha Nussbaum belong,[9] have raised serious objections to the traditional conception of moral theory which also need to be taken seriously in computer ethics. Anti-theorists assert that it is not the case that all correct moral judgements and practices can be deduced from universal, timeless principles, which it is the job of moral theory to articulate; that all moral values are commensurable on a common scale which it is the task of moral theory to provide; that all disagreements and conflict can be solved by means of the application of theory and the use of a decision procedure which it is the job of moral theory to supply; that moral theory is entirely normative.

According to this approach, there may be points of diminishing returns of moral theorizing in computer ethics, since, as Nagel has pointed out, 'our capacity to resolve conflicts in particular cases may extend beyond our capacity to enunciate general principles that explain those solutions'.[10] According to Nagel, 'to look for a single general theory of how to decide the right thing to do is like looking for a single theory of how to decide what to believe'.[11]

The other line of anti-theoretical critique concerning ethical theory is not so much that ethical theory is impossible, unnecessary or undesirable, but that it is useless in practice, except perhaps as reminder of the importance of a particular type of value and value-based arguments or as a summary of past experiences. Richard Posner (1999), who sympathizes with the anti-theoretical position, lodged an attack on the usefulness of standard moral theory to which he refers as *Academic Moralism*, i.e. the assumption that ethical theory as studied at universities in philosophy departments around the world can help us to arrive at better understandings and solutions of our practical problems. Posner, with his long experience as a judge (Chief Judge of the US Court of Appeals of the Seventh Circuit) and with a thorough knowledge of academic ethics, denies that moral theory is at all useful in practice, and that it has any policy impact. Academic moralism is not an agent of moral change. According to Posner it fails as an agent of moral change partly because those who work on it fail to make it so: 'Unhindered by external checks and balances, the academic moralist has no incentive to be useful to anybody ... the intellectual gifts moral philosophers exhibit need not, and in their normative work usually do not, generate a positive social product' (p. 80). Posner's critique is coarse-grained, but not unfounded, and touches a delicate open nerve of modern practical philosophy. Moral theory

---

[9] See for a collection of essays in Anti-theory Clarke and Simpson 1989.

[10] Thomas Nagel 'The Fragmentation of Value', reprinted in Gowans 1987, pp. 174–187, p. 181.

[11] Thomas Nagel 'The Fragmentation of Value', reprinted in Gowans 1987, pp. 174–187, p. 181.

as it stands now is only marginally relevant to the world where the all-important decisions are made. It may eventually become obsolete if it does not deliver on its constitutive promise to be relevant to practice and the professions.

The upshot of this characterization of the starting points for moral theory and computer ethics is that different types of value (utility and outcomes, general rights and principles, specific obligations, agent's commitment to own projects and perfectionist ends) can always be brought to bear upon morally problematic situations, sometimes in the form of free-standing considerations which have to be balanced against others, sometimes in the form of applications of a general principle (e.g. principle of utility, categorical imperative) to a particular case. Contributions to thinking about the hard questions of ICTs hardly ever present themselves in the form of elaborate and thorough applications of austere Aristotelian, Kantian or utilitarian theories. In a sense, they are superfluous as theories, but not as sources of moral arguments and moral considerations. Enlightening contributions in computer ethics[12] use arguments, insights and considerations which are inspired and informed by Kantian, utilitarian or Aristotelian views and by the values that are central to them.

4.3.3   Methodology

With respect to 'problematic situations', as Dewey called them – whether that is at the individual, professional, institutional or societal level – we thus need to make up our minds and come to a conclusion in the midst of a panoply of considerations. Since there is no standard method or decision procedure to solve conflicts between different types of values, and unify our thoughts, we will have to do with the ancient, but notoriously elusive, resource of 'practical wisdom', i.e. the weighing, sizing up the situation, seeing what is morally salient and relevant to achieving one's moral goals and choosing the appropriate course of action.

There are not many methodological constraints to ethical thinking apart from (1) an epistemic obligation to explain why one holds certain moral beliefs and not others, (2) to do all that is in our power to free our actions from the defects of ignorance, error and possible bias, and (3) to eliminate inconsistencies in our moral belief set by applying the logic of moral reasoning, which (4) typically comprises the application of the principle of supervenience of moral reasons. The Principle of Supervenience states that there are no moral differences without differences in other non-moral respects.[13] To use the example provided by Richard Hare, on the basis of which the notion of

---

[12] See footnote 1 for an overview.
[13] See *Stanford Encyclopedia of Philosophy* online, article on supervenience.

supervenience gained currency in Ethics: to state that 'room 13 is a nice room, but room 12, although similar in all relevant respects, is not a nice room' is to make a self-contradictory statement (Hare 1984).

A further general methodical directive concerns the contemporary orthodoxy about 'the way we do ethics now' as James Griffin (1993) has called it. It is neither a decision procedure, nor a method in a strict sense, but a sketch of a way of proceeding used by all sensible people who have access to relevant facts of the matter, have moral values, elementary logic, and ideals of clarity and consistency. It is, in essence, a coherence model along the lines of the method of Wide Reflective Equilibrium (Griffin 1996, van den Hoven 1997), which occupies the middle ground between generalist and particularist views, between theoretical and anti-theoretical construals of moral thinking, but which retains the principle of supervenient application of moral reasons – or the universalizability of moral considerations – as a requirement of rationality in public moral discourse concerning practical problems.

For computer ethics, neither the simple engineering view of application along the lines of the deductive nomological model of explanation in physics (or for example the simple practical syllogism) nor the opposite extreme of particularism seems viable. Coherence models of moral justification allow for the desired level of logical structure and generality in our moral belief sets without becoming impervious to the force of contextual and agent-relative considerations.

What we need in applied ethics of computing is what Nagel describes as a 'method of breaking up or analyzing practical problems to say what evaluative principles apply, and how'. This method 'would simply indicate the points at which different kinds of ethical considerations needed to be introduced to supply the basis for a responsible and intelligent decision'.[14]

4.3.4   Applying moral theories

Aristotelian ethics answers questions about what to do on the basis of what virtue requires or what a virtuous person would do. A person is virtuous when he has moral virtue(s), i.e. character traits or dispositional properties, which allow him to choose and act in order to achieve happiness or human flourishing. Moral virtues, such as modesty, courage, justice, are learned by following moral exemplars. Ideally, the virtuous person also possesses a general intellectual capacity, practical wisdom – which enables him to identify the morally relevant features in every situation and determine the right course of action. Moral knowledge, which is thus embedded in a person's character, is motivational, i.e. it is impossible to know what is right and not be inclined to do it.

---

[14] Thomas Nagel 'The Fragmentation of Value', reprinted in Gowans 1987, p. 184.

Utilitarian moral theories instruct one to choose those actions which have the best consequences or outcomes. More specifically, they require one to choose those actions – or to choose those rules or policies for acting – that bring about the greatest good for the greatest number in the world. The good is measured in some quantity of non-moral good such as happiness, pleasure, well-being or money. When confronted with a choice between different courses of action, one ought to choose that course of action which maximizes utility compared to the alternatives. There are several versions of utilitarianism to which we cannot do justice here, but their overall structure is the same. There is some end which is good by independent non-moral criteria and which is brought about by means of the agent's actions or indirectly by the rule which is followed in action. This means–end relationship and causal relationship (Nozick (1993) allows this relation also to be symbolic apart from causal) confers moral status on the action or the state. *The right* is thus defined in terms of *the good* in utilitarian theories.

Kantian theories state that whether an action is obligatory does not depend on its consequences, but on characteristics of the action itself and its compliance with the highest ethical principle: the Categorical Imperative. The most accessible formulation of the categorical imperative states that one ought to respect human beings as such and not use them as mere instruments for one's purposes. According to an alternative formulation, one ought to choose that course of action which instantiates a policy that can without contradiction be adopted by everyone or that can be willed to be a universal law. Kantian accounts are, in a sense, the antidote to utilitarian theories. Each human being is a source of meaning and value, has a life of his own, is morally autonomous and deserves to be respected as such, whatever the consequences. Both utilitarianism and Kantian moral theories are universalist and agent-neutral. Utilitarians apply the criterion for moral standing (sentience) universally and Kantians apply their criterion of moral standing (rationality) to all (and only) rational beings (including angels and artificial intelligences).

As an illustration of how these different normative ethical theories can figure in debates about new and emerging ICTs issues, we will look at how they figure in the discussion about ultra violent computer games (Wonderly 2008, Waddington 2007). In ultra violent computer games such as *Grand Theft Auto*, *Zog's Nightmare* and *Manhunt*, players are invited to run extermination camps, kill for snuff movies, and run over people to score points. Parents who watch their children play these games may have moral concerns and many others would understand their concerns. Is there anything morally wrong with playing ultra violent computer games and, if so, what is it?

Utilitarian accounts seem to fail to account for the concerns, since there are no relevant other moral entities (sentient creatures) harmed by the action of those who play violent computer games. There is only virtual suffering and

virtual pain. Mill thought that even harmless acts could be morally forbidden if they violated good manners or gave offence. Clearly, if a group community of individuals created a violent computer game and played it among themselves, without anyone knowing about it, there could be no offence or violation of good manners and, consequently, no indirect harm. Alternatively, one could say that this pastime is bound to affect someone's behaviour towards his fellow human beings and it is likely to bring about negative effects. The problem with this suggestion is that there is no conclusive evidence that it would. Television has been around for almost half a century and still the debate over whether violent movies trigger violent behaviour continues. If there were some remarkable statistical evidence of this effect, this would not show that there is a causal connection between playing and violent behaviour. Again if the statistical evidence combined with psychological and neurological evidence proved the nexus for a small percentage of the population beyond reasonable doubt, the question would still remain whether the dis-utilities (some occasional violent behaviour) outweighed the positive utilities (long happy hours of gaming for the millions).

Kantian accounts fare no better. There are no rational human beings affected by this sort of game playing, apart from the player himself. No one is used as a mere instrument and no one's dignity is at stake, except the dignity perhaps of the player himself. It is even possible to imagine that everyone engages in solitary violent computer gaming, without contradicting oneself in the relevant sense. It also seems possible to subscribe to a universal law which says that everyone should spend some time every day playing ultra violent computer games, although that may sound a bit awkward.

One could stretch the Kantian view by using Kant's argument against the cruelty against animals. Kant was not so much concerned with animals as such. They are not rational beings in the relevant sense, so they do not qualify for moral standing. But he was opposed to cruelty against animals because he believed that this type of behaviour corrodes one's character and is likely to facilitate cruelty against human beings (Midgley 1985, Brey 1999). Likewise, cruelty against virtual humans could predispose to cruelty against real humans. Another way of applying Kantian Ethics is by construing the playing of violent computer games as the violation of an (imperfect) duty to oneself. According to Kant, every human being has a duty to himself to cultivate his capacities, his moral and non-moral capacities and talents (Dennis 1997). In virtually killing, going through the motions, rehearsing, engaging in role-playing, without any artistic or educational idea, one is not respecting humanity as it is exemplified in one's own personality.

What seems objectionable in playing violent computer games is the thought that a person is spending a considerable amount of time identifying with a character in a game who is in mental states which are relevantly similar to those involved in offline killing, rapings and torturing, and gets rewarded

for it by scoring points. McCormick (2001) argues that neither utilitarian nor Kantian accounts can demonstrate the moral wrongness of these proceedings. He suggests that only an Aristotelian account can explain our moral intuitions concerning them: 'by participating in simulations of excessive, indulgent, and wrongful acts, we are cultivating the wrong sort of character'. Wonderly (2008) – with reference to Hume – claims that empathy has a central role in making moral judgements and that research shows that playing violent video games is inimical to the fostering of empathic functioning.

It is difficult to account for our moral apprehension and to locate the moral wrongness on the basis of the values of utility or general rights. The Humean and Aristotelian approaches seem most promising, but if on the other hand conclusive evidence would become available to show that gamers are indeed inclined to be violent in the real world we would probably stop worrying about their moral characters and turn to a straightforward utilitarian account to justify our concern with this type of application.

4.4   Mid-level theories: ground preparation and conceptual reconstructions

Instead of applying highly abstract traditional ethical theories straightforwardly to particular ICTs issues, it is often more helpful to utilize mid-level normative ethical theories, which are less abstract, more testable and which focus on technology, interactions between people, organizations and institutions. Examples of mid-level ethical theories are Rawls' theory of justice, which could be construed as broadly Kantian, Amartya Sen and Martha Nussbaum's capability approach, which can be construed as broadly Aristotelian, and Posner's economic theory of law, which is broadly utilitarian. These theories already address a specific set of moral questions in their social, psychological, economic or social context. They also point to the empirical research that needs to be done in order to apply the theory sensibly. I have elsewhere discussed how these mid-level theories may be fruitfully applied (van den Hoven 2005, 2008 van den Hoven and Rooksby 2008) and more work is done on them to make them even better suited for application to real-life problems. Concerning violent computer games, for example, the capability approach of Martha Nussbaum seems to capture what concerns parents and those who sympathize with them. Coeckelbergh (2007) uses Martha Nussbaum's capability approach to argue that the trained insensitivity towards human suffering, which goes on in playing violent computer games, is inimical to cultivating humanity and squarely opposed to training the fine sensibility and awareness required for moral excellence and human flourishing. Nissenbaum and others have started to work on how games may

be designed that build in moral desirable features and capability enhancing elements.[15]

Floridi's Information Ethics[16] provides a high-level value theory which applies to the ICTs domain, which at the same time allows for specification at the mid-level and lower levels of abstraction and specification. It is universally applicable also outside the ICTs domain in a stricter sense, and construes information as ontologically fundamental and entropy – in the specific sense of destruction, damage and vandalizing of informational entities and environments – as the morally most relevant category. According to information ethics along these lines the moral status of actions concerns their informational status and information objects thus have moral significance and are hence deserving of respect (Sicart and Studies 2009). Computer ethics should thus be concerned with finding out what increases entropy and which actions and events counteract it. Information Ethics is a recent alternative to traditional ethical theory to account for the moral phenomena and is the subject of further research to investigate how it can be made to bear upon the practical problems in ICTs[17] and to demonstrate that it has an explanatory and justificatory surplus compared to the traditional ethical normative theories. Relating to the topic discussed, Miguel Sicart has applied Floridi's Information Ethics to tackle problems in the design of computer games (Sicart 2009).

Important for the application of the range of mid-level ethical theories, which are specifications (Moor 1985) of high-level ethical theories, is what Bernard Gert and Carl Danner Clouser have called 'ground preparation', i.e. the meticulous understanding of the field to which ethical theory is being applied. This is part of ethics itself and it may well be considered as the essence of applied ethics. It requires more analytical skills and rigour, according to Clouser, than is generally thought (Clouser 1980). We need to know what the properties of artificially intelligent agents are, how they differ from human agents; we need to establish what the meaning and scope is of the notion of 'personal data', what the morally relevant properties of virtual reality are. These are all examples of preparing the ground conceptually before we can start to apply normative ethical considerations.

Jim Moor has suggested that, with respect to many issues in computer ethics, we are confronted with a conceptual vacuum and an ensuing policy vacuum in these and other cases (van den Hoven 2005). I suggest, in addition, that we are also confronted with a *design vacuum* and we are at a loss which systems to make, which software to engineer, which lines of computer code to write. Therefore, an important part of the ground preparation consists in conceptual reconstruction of the key concepts before any values, principles or theories

---

[15] Flanagan, Howe and Nissenbaum in van den Hoven and Weckert 2008, pp. 322–354.

[16] See special issue on Floridi of *Ethics and Information Technology*, vol 10, 2008, nos 2–3.

[17] Richardson (1990) defended the model of norm specification as covering a middle ground position between deduction and balancing.

can be applied. Reconstruction is a process of articulating and formulating specific adequate conceptions of general notions (and articulating criteria of adequacy) that have become problematic in their application to a world that has changed since the time these notions gained currency.

John Rawls made a distinction between *concepts* and *conceptions* of justice (Rawls 1971) which is pertinent in this context. Many people share the general *concept* of justice (or equality or responsibility for that matter), without necessarily sharing the same *conception* of justice. Conceptions are the specific and substantive specifications and instantiations of a general and formal concept. Rawls famously proposed his conception of *justice as fairness*, but utilitarians have proposed radically different conceptions of justice. Our philosophical notions are 'essentially contested concepts' as W. B. Gallie has called them (Gallie 1956). Controversy over the correct meaning – or discussion of the most adequate conceptions which ought to be construed as the action guiding instantiation of them – has become part and parcel of their meaning. ICTs prompt us to revisit traditional *conceptions* of privacy, responsibility, property, democracy, community and formulate new and more appropriate or interesting conceptions, which serve and suit us better. Dewey defined this reconstruction as one of the main tasks of philosophy and he saw it as a process that never stops. In a rapidly changing world, traditional conceptions are like tools that have deteriorated in use and therefore need to be maintained and reconstructed in order to keep them fit for the task at hand.

Discussions in computer ethics are about '*digital* democracy', '*software* patents', '*virtual* child pornography', '*online* relationships', '*net* friendship', '*cyber* communities', '*informational* privacy', '*artificial* life', '*tele*-work', '*intellectual* property', '*e*-Trust' and '*electronic* Commerce'. This semantic expansion – the result of adding qualifications in the form of prefixes (cyber, virtual, digital, informational, e, electronic, tele, software) from the ICT domain to traditional concepts – may also suggest that, since we have the fancy terminology, we also have come to grips with the phenomena which are conjured up by the new techno-speak and that we know what to do in terms of design, policy and law. But as Moor correctly suggested, this is often not the case: what is a 'net-friend', 'e-Trust', etc.?

The concept of democracy is widely used all over the world in different historical periods to indicate some sort of involvement of the people in the political process. Governments all over the world are now investing considerable amounts of taxpayers' money in online democracy. Which conception of democracy are they using? There have been radically different, substantive *conceptions* of democracy. One may have a so-called *direct conception of democracy*, or a *deliberative*, or *participative*, or *representative* conception. Different conceptions have quite different technologies to support or express them. Direct Democracy ICT projects would heavily invest in online voting technology; deliberative and participatory conceptions point in the direction

of projects which aim at establishing forms of deliberation, discussion and sharing ideas between citizens online, which requires a completely different set of technologies. Pettit's conception of *contestatory* democracy[18] would point in the direction of checks and balances and tools for citizens to get access to relevant information and effectively protest and contest government decisions online.

The fact that we talk about *cyber communities* does not imply that we actually understand the nature of communities any better than we did before, let alone that we have a clear idea about sociality, community and individuality online, that we know whether to regulate them and how, understand what their value is in individual identity formation, what levels of security should be offered and whose responsibility it is. What we seem to be saying, when we use the term, is that we do not yet know exactly what we mean, but that it has something to do with people getting together, interacting, getting to know each other, exchanging information, embarking upon coordinated and joint action, identifying with common goals, and that they do all this online, without having to meet face to face. In talking about 'cyber community' we are taking out a mortgage on a future analysis and conceptual reconstruction of a conception of 'cyber community'.

This would not be a special problem if we did not have to draft policies, laws and regulation and design information systems and program computers on a daily basis, if we did not have to proceed in practice. But we do. The design of procedures, institutions, systems, information architectures and computational devices requires articulation, precision and detail. It requires precision in the formulation of our ideas and the specification of what we want to achieve by means of the technology.

4.5   Design

Moral problems in professional ethics literature often take the form of a moral dilemma. A professional in a dilemmatic situation has at least two obligations, but he cannot fulfil both of them at the same time. What should the professional do? One type of reaction to dilemmatic situations is to make the best of them and to try to see how one can limit the damage – one might engage in utilitarian calculations, in Kantian reflections or ask what a virtuous person would do in that situation to find how to make the best of it. Moral thinking about such dilemmas assumes that *the situation is given.* In a straightforward sense that is a correct construal because it is often a thought experiment, but what this mode of moral thinking and theorizing about these dilemmatic thought experiments is suppressing is the fact that the problematic situations in real life, which constitute moral dilemmas, are the

---

[18] See for a discussion van den Hoven 2005.

result of hundreds of prior design decisions and choices. This may be illustrated by reference to one of the most discussed dilemmatic thought experiments in contemporary ethics, the Trolley Case. Suppose you are at the forking path of a downhill railway track and a trolley cart is hurtling down and will pass the junction where you stand. There is a lever which you can operate. If you do nothing the trolley will kill five people, who are tied down to the track further downhill. If you pull the lever, the trolley will be diverted to the other track where there is only one person tied to the track. Is it morally permissible to pull the lever, or is there even a moral obligation to do so?

Engineers and other sensible non-philosophers often reply to Trolley Cases by saying that it is a stupid piece of infrastructure that should have been designed differently. This is not a proper move in the philosophy language game, but is a most interesting move in another language game, namely the one we adopt when we talk about preventing deaths, avoiding tragic moral choices and improving the world in the future. The obsession with moral theory and its refinement blinds one to an important aspect of moral thinking, namely *design*.

Especially those with a technology and engineering background may be able to suggest all kinds of clever design solutions that would prevent this tragic situation from occurring in the future. Their natural attitude to the problems as presented is to formulate solutions to real-world problems instead of contributing to refining ethical theories on the basis of crude and information-poor examples.

Moral analysis of the situation needs to deal with the history of choices and design and development antecedents. Computer ethicists should therefore probe beyond the status quo and ask how the problem came into being and what the design and architectural decisions are that have led up to it. We will not be able to resolve Trolley problems to our full satisfaction once they have presented themselves to us. We need to try to prevent them from occurring in the first place. As Ruth Barcan Marcus has stated, we have a higher-order obligation or a higher-order responsibility (Barcan Marcus in Gowans 1987, p. 200) to prevent situations in which we ourselves and others cannot meet their responsibility and do what ought to be done, 'One ought to act in such a way, that if one ought to do X and one ought to do Y, then one can do both X and Y.'

Cass Sunstein has pointed out (Thaler and Sunstein 2008) that most professionals – ICTs architects and ICT professionals are eminent examples in this respect – are *choice architects*, who 'have responsibilities for organizing the context in which people make decisions'. They design either for tragic choices and likely accidents or for responsibility and safety and security.

As far as the institutional dimensions of moral situations are concerned, this design type of question is now being addressed more often. The question is now posed, which institutional and material conditions need to be fulfilled

if (1) we want to prevent situations where the best we can do is limit the damage and (2) we want the results of our ethical analyses to be implemented? How can we increase the chances of changing the world in the direction in which our moral beliefs – held in wide reflective equilibrium – point? How can we design the systems, institutions, infrastructures and ICTs applications in the context of which users will be able to do what they ought to do and which will enable them to prevent what ought to be prevented (Turilli 2007, Turilli 2008)? I have dubbed this notable shift in perspective 'The Design Turn in Applied Ethics' (van den Hoven 2008c, van den Hoven, Miller and Pogge 2010). The work of John Rawls for the first time gave rise to talk about design in ethics. Thinking about social justice can, in the context of Rawls' theory, be described as formulating and justifying the principles of justice in accordance with which we should design the basic institutions in society. Thomas Pogge, Russell Hardin, Cass Sunstein, Robert Goodin, Dennis Thompson and others (van den Hoven 2008) have taken moral theory and applied ethics a step further down this path of semantic descent and practicality. Not only do they want to offer applied ethical analyses, they also want to think about the economic conditions, institutional and legal frameworks and incentive structures that need to be realized if our applied analyses are to stand a chance in their implementation and thus contribute to bringing about real and desirable moral changes in the real world. Design in the work of these authors is primarily focused on institutional design, but the Design Turn clearly brings into view the design of socio-technical systems, technological artefacts and socio-technical systems. This suggests in part another mode of moral thinking.

To sum up: high-level, moral theories – which each put different types of moral value centre stage – are to be specified and exemplified in a process of clarification of the moral issues of information societies in the form of mid-level theories. Mid-level theories may then in turn be used as sources of moral arguments in the relevant empirical domains, where conceptual reconstructions have prepared the ground for their application. Reconstructed concepts, e.g. *contestatory democracy, justice as fairness, privacy as data protection*, function as high-level architectural principles for the design of information systems and ICTs applications. These principles can be utilized as non-functional requirements, which can be further specified by means of functional decomposition in specifications for the development of ICTs applications.